

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 04:17:29 ; Search time 2848.01 Seconds  
(without alignments)  
9405.294 Million cell updates/sec

Title: US-09-864-675-3  
Perfect score: 897  
Sequence: 1 atgaggcgcgacccggcccc.....caatggtcaacttctcctaa 897

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*

28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	8						Description
	No.	Score	Query Match	Length	DB	ID	
	1	674	75.1	805	12	BI918620	BI918620 603176570
	2	565.6	63.1	1047	12	BM914622	BM914622 AGENCOURT
c	3	467.2	52.1	1041	12	BI412864	BI412864 602988202
	4	467	52.1	524	13	BX281777	BX281777 BX281777
	5	408.2	45.5	549	9	AA706226	AA706226 ah28a07.s
	6	396.6	44.2	412	9	AI041451	AI041451 ow36c02.s
	7	363.8	40.6	427	10	BF108794	BF108794 7152g03.x
	8	318.6	35.5	488	13	BX529505	BX529505 BX529505
c	9	255	28.4	949	12	BI410828	BI410828 602963734
	10	234.4	26.1	333	10	BE983573	BE983573 UI-M-CG0p
	11	224	25.0	297	9	AA772412	AA772412 ai44e12.s
	12	210.2	23.4	795	12	BI651936	BI651936 603298677
	13	195.2	21.8	259	10	BE648780	BE648780 UI-M-BH2.
	14	195.2	21.8	327	9	AA968077	AA968077 uh09h01.r
c	15	182.6	20.4	362	13	BX089049	BX089049 BX089049
	16	169.8	18.9	529	10	AW476657	AW476657 uq79e01.y
c	17	154.8	17.3	539	9	AL918370	AL918370 AL918370
	18	142	15.8	657	13	BQ078813	BQ078813 fy81c06.y
	19	111	12.4	256	10	AW762061	AW762061 ur53c01.y
	20	107.6	12.0	458	9	AI152190	AI152190 ud18h10.r
	21	105.2	11.7	750	28	BZ847665	BZ847665 CH240_239
	22	101	11.3	493	28	BH057870	BH057870 RPCI-24-9
c	23	100.4	11.2	481	28	AZ987593	AZ987593 2M0270P10
c	24	97.4	10.9	180	29	CE103297	CE103297 tigr-gss-
	25	90.2	10.1	243	10	BB570162	BB570162 BB570162
c	26	85	9.5	512	9	AI073386	AI073386 ool3d06.x
	27	81.2	9.1	167	9	AI836531	AI836531 UI-M-AP0-
	28	76.6	8.5	477	10	BE984041	BE984041 UI-M-CG0p
	29	70.8	7.9	761	13	BX876483	BX876483 BX876483
	30	67.2	7.5	765	12	BI522417	BI522417 603175321
	31	67	7.5	769	12	BI413085	BI413085 602990205
	32	66	7.4	751	29	CNS04J6G	AL293137 Tetraodon
	33	64.8	7.2	321	10	BE983721	BE983721 UI-M-CG0p
	34	64.4	7.2	538	9	AL925790	AL925790 AL925790
c	35	61	6.8	491	9	AL909688	AL909688 AL909688
c	36	60.2	6.7	322	9	AL909689	AL909689 AL909689
	37	59	6.6	356	29	CG614660	CG614660 OST303096
	38	58.8	6.6	363	29	CE326275	CE326275 tigr-gss-
	39	58.8	6.6	413	14	N62228	N62228 yz63c08.s1
	40	58.4	6.5	685	14	CA351220	CA351220 622234 NC
	41	53	5.9	1630	11	AK051824	AK051824 Mus muscu
	42	51.6	5.8	925	29	CNS0091P	AL053013 Drosophil
c	43	51.6	5.8	982	13	BX415111	BX415111 BX415111
	44	50.8	5.7	647	12	BI960178	BI960178 HVSMEn002
	45	50.4	5.6	251	9	AW045376	AW045376 UI-M-BH1-

## ALIGNMENTS

## RESULT 1

BI918620

LOCUS BI918620 805 bp mRNA linear EST 16-OCT-2001

DEFINITION 603176570F1 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5240969 5', mRNA sequence.

ACCESSION BI918620

VERSION BI918620.1 GI:16182295

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 805)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11607 row: k column: 18

High quality sequence start: 2

High quality sequence stop: 778.

## FEATURES

source

Location/Qualifiers

1. .805

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5240969"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_121"

/note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 75.1%; Score 674; DB 12; Length 805;

Best Local Similarity 98.7%; Pred. No. 2.8e-127;

Matches 732; Conservative 0; Mismatches 5; Indels 5; Gaps 5;

Qy 1 ATGAGGCGCGACCCGGCCCCCGGC-TTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTG 59  
 |||  
 Db 64 ATGAGGCGCGACCCGGCCCCCGGCGTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTG 123  
 |||  
 Qy 60 CTACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGGA 119  
 |||  
 Db 124 CTACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGGA 183  
 |||  
 Qy 120 GGGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCC 179  
 |||  
 Db 184 GGGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCC 243  
 |||  
 Qy 180 GCCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGG 239  
 |||  
 Db 244 GCCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGG 303  
 |||  
 Qy 240 GGGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCA 299  
 |||  
 Db 304 GGGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCA 363  
 |||  
 Qy 300 GCGCTACATCTTTTTCTGGAGCCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTG- 358  
 |||  
 Db 364 GCGCTACATCTTTTTCTGGAGCCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGT 423  
 |||  
 Qy 359 CCCCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTG 418  
 |||  
 Db 424 CCCCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTG 483  
 |||  
 Qy 419 ACTGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGA 478  
 |||  
 Db 484 ACTGCGCCACCCGGCCCCAAGTTGAAGAAGATGACGAGCCAGACGGGACAGGTGGGTGAGA 543  
 |||  
 Qy 479 AGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCA 538  
 |||  
 Db 544 AGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCA 603  
 |||  
 Qy 539 AGGATGGCAAGGAGCTCAACCG-CAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGA 597  
 |||  
 Db 604 AGGATGGCAAGGAGCTCAACCGTCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGA 663  
 |||  
 Qy 598 AAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGA-GTATGTCTG 656  
 |||  
 Db 664 AAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGGTATGTCTG 723  
 |||  
 Qy 657 CGAGG-CCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCG 715  
 |||  
 Db 724 CGAGGCCCAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGTTTTACGTCAACAGGT 783  
 |||  
 Qy 716 TGAGCACCACCCTGTCATCCTG 737  
 |||  
 Db 784 TGAGCACCAACCTGTCATCCTG 805

RESULT 2

BM914622

LOCUS BM914622 1047 bp mRNA linear EST 12-MAR-2002

DEFINITION AGENCOURT\_6615334 NIH\_MGC\_113 Homo sapiens cDNA clone IMAGE:5480308

5', mRNA sequence.

ACCESSION BM914622

VERSION BM914622.1 GI:19365001

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1047)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Dr. Mark Watson  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM2002 row: p column: 05  
High quality sequence stop: 541.

FEATURES  
source Location/Qualifiers  
1. .1047  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5480308"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_113"  
/note="Organ: spleen; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCACGAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

#### ORIGIN

Query Match 63.1%; Score 565.6; DB 12; Length 1047;  
Best Local Similarity 96.8%; Pred. No. 4.4e-105;  
Matches 577; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy	272	GCTCCTGTGTGCCGCTCGAAAGGAACCAGCGCTACATCTTTTTCCTGGAGCCCACGGAAC	331
Db	1	GCTCCTGTGTGCCGCTCGAAAGGAACCAGCGCTACATCTTTTTCCTGGAGCCCACGGAAC	60
Qy	332	AGCCCTTAGTCTTTAAGACGGCCTTTGCCCCCTCGATACCAACGGCAAAAATCTCAAGA	391
Db	61	AGCCCTTAGTCTTTAAGACGGCCTTTGCCCCCTCGATACCAACGGCAAAAATCTCAAGA	120
Qy	392	AAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGTTGAAGAAGATGA	451
Db	121	AAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGTTGAAGAAGATGA	180

QY 452 AGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGTA 511  
 |||  
 Db 181 AGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGTA 240  
 QY 512 ATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACA 571  
 |||  
 Db 241 ATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACA 300  
 QY 572 TTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGG 631  
 |||  
 Db 301 TTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGG 360  
 QY 632 TGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCC 691  
 |||  
 Db 361 TGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCC 420  
 QY 692 GGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCATCCTGGTCGGGGCAGCCCC 751  
 |||  
 Db 421 GGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCATCCTGGTCGGGGCAGCCCC 480  
 QY 752 GGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCG 811  
 |||  
 Db 481 GGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGCCTGCTACTACATCG 540  
 QY 812 AGGGCATCAACCAGCTCTCCTGCAAAGTGTCCTGTGGGATACACCGGGGACAGGTGT 867  
 |||  
 Db 541 AGGCCATCAATCAGCTTTCTGCAAATGTCCCAATGGATTCTTCCGACCAACATGT 596

# RESULT 3

BI412864/c

LOCUS BI412864 1041 bp mRNA linear EST 14-AUG-2001  
 DEFINITION 602988202F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5144016 5',  
 mRNA sequence.

ACCESSION BI412864

VERSION BI412864.1 GI:15173787

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1041)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11355 row: d column: 01

High quality sequence start: 11



```
Db      236 TCAAGTATGGCAATGGCAGAAAGAACTCACGGCTACAGTTCAACAAAGTGAGGGTGGAGG 177
QY      638 ACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGGCC 697
        | ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      176 ATGCCGGGGAGTACGTCTGTGAGGCCGAGAACATCCTTGGGAAGGACACCGTGAGGGGCC 117
QY      698 GGCTTTACGTCAACAGCGTGAGCACACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGT 757
        | || | ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      116 GACTCCATGTCAACAGCGTGAGCACCACTCTGTATCCTGGTCGGGCATGCCCGGAAGT 57
QY      758 GCAACGAGACAGCCAAGTCCTA--TTGCGTCAATGGAGGCGTCTGCTACT 805
        |||| ||||| ||||| ||||| || ||| ||||| |||||
Db      56 GCAATGAGACCGCCAAGTCCTACCATGTGTGAATGGAGGCGTGTGCTACT 7
```

```
FEATURES
    source
        Location/Qualifiers
            1. .524
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGp998K1811607 ; IMAGE:5240969"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_121"
                /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
```



Site\_2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH\_MGC Library."

# ORIGIN

Query Match 52.1%; Score 467; DB 13; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-85;  
 Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60
          |||
Db      58 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 117

Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
          |||
Db     118 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 177

Qy     121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
          |||
Db     178 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 237

Qy     181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
          |||
Db     238 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 297

Qy     241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
          |||
Db     298 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 357

Qy     301 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
          |||
Db     358 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 417

Qy     361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
          |||
Db     418 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 477

Qy     421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACA 467
          |||
Db     478 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACA 524
  
```

## RESULT 5

AA706226

LOCUS AA706226 549 bp mRNA linear EST 12-JAN-1999  
 DEFINITION ah28a07.s1 Soares\_parathyroid tumor NbHPA Homo sapiens cDNA clone  
 1240116 3' similar to TR:P43328 P43328 NEU DIFFERENTIATION FACTOR  
 NDF04 ;, mRNA sequence.

ACCESSION AA706226

VERSION AA706226.1 GI:2716144

KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 549)  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

Possible reversed clone: similarity on wrong strand  
 Possible reversed clone: polyT not found  
 Insert Length: 689 Std Error: 0.00  
 Seq primer: -40m13 fwd. ET from Amersham  
 High quality sequence stop: 451.

FEATURES  
 source Location/Qualifiers  
 1. .549  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="1240116"  
 /tissue\_type="parathyroid tumor"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares\_parathyroid\_tumor\_NbHPA"  
 /note="Organ: parathyroid gland; Vector: pT7T3D  
 (Pharmacia) with a modified polylinker; Site\_1: Not I;  
 Site\_2: Eco RI; 1st strand cDNA was primed with a Not I -  
 oligo(dT) primer  
 [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGACCAATTTTTTTTTTTTTTTTTTTT  
 TTTT-3'], double-stranded cDNA was size selected, ligated  
 to Eco RI adapters (Pharmacia), digested with Not I and  
 cloned into the Not I and Eco RI sites of a modified pT7T3  
 vector (Pharmacia). Library went through one round of  
 normalization to a Cot = 5. Library constructed by Bento  
 Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid  
 adenomas was kindly provided by Dr. Stephen Marx, National  
 Institute of Diabetes and Digestive and Kidney Diseases,  
 NIH."

#### ORIGIN

Query Match 45.5%; Score 408.2; DB 9; Length 549;  
 Best Local Similarity 91.7%; Pred. No. 5e-73;  
 Matches 431; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 424 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 483  
 |||||  
 Db 15 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 74

QY	484	TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT	543
Db	75	TCGCTGAAGTGTGAGGCAGCAGCGGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT	134
QY	544	GGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAGAAC	603
Db	135	GGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAGAAC	194
QY	604	TCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCC	663
Db	195	TCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCC	254
QY	664	GAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACC	723
Db	255	GAGAACATCCTGGGGAAGGACACCGTCCGGAGGCCGGCTTTACGTCAACAGCGTGACGACC	314
QY	724	ACCCTGTCATCCTGGTTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGC	783
Db	315	ACCCTGTCATCCTGGTTCGGGGCACGCCCGGAAGTGCAACGNGACAGCCAAGTCCTATTGC	374
QY	784	GTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCT	843
Db	375	GTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGGCACCT	434
QY	844	GTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGGTCAACTTCTC	893
Db	435	GGGCTGCACTGCTTAGAACTTGGTACCCAGAGCCACCACTTCCCCATCTC	484

Trace considered overall poor quality  
 Insert Length: 671 Std Error: 0.00  
 Seq primer: -40ml3 fwd. ET from Amersham  
 High quality sequence stop: 1.

# FEATURES

source

Location/Qualifiers

1. .412

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:1648898"

/tissue\_type="parathyroid tumor"

/dev\_stage="adult"

/lab\_host="DH10B (ampicillin resistant)"

/clone\_lib="Soares\_parathyroid\_tumor\_NbHPA"

/note="Organ: parathyroid gland; Vector: pT7T3D

(Pharmacia) with a modified polylinker; Site\_1: Not I;

Site\_2: Eco RI; 1st strand cDNA was primed with a Not I -  
 oligo(dT) primer

[5'-TGTTACCAATCTGAAGTGGGAGCGGCCGACCAATTTTTTTTTTTTTTTTTTTT  
 TTTT-3'], double-stranded cDNA was size selected, ligated  
 to Eco RI adapters (Pharmacia), digested with Not I and  
 cloned into the Not I and Eco RI sites of a modified pT7T3  
 vector (Pharmacia). Library went through one round of  
 normalization to a Cot = 5. Library constructed by Bento  
 Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid  
 adenomas was kindly provided by Dr. Stephen Marx, National  
 Institute of Diabetes and Digestive and Kidney Diseases,  
 NIH."

## ORIGIN

Query Match 44.2%; Score 396.6; DB 9; Length 412;

Best Local Similarity 97.6%; Pred. No. 1e-70;

Matches 402; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy	426	CACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATC	485
Db	1	CACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATC	60
Qy	486	GCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAGGATGG	545
Db	61	GCTGAAGTGTGAGGCAGCAGCGATAAATCCCCAGCCTTCCTACCGTTGGTTCAGGATGG	120
Qy	546	CAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAACTC	605
Db	121	CAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAACTC	180
Qy	606	ACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGA	665
Db	181	ACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGA	240
Qy	666	GAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCAC	725
Db	241	GAACATCCTGGGGAAGGACACCGTACGAGGCCGGCTTTACGTCAACAGCGTGAGCACCAC	300
Qy	726	CCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGCGT	785

Db 301 CCTGTCATCCTGGTCGGGGCACGCCGGGAAGTGCAACGNGACAGCCAAGTCCTATTGCGT 360

QY 786 CAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAG 837  
 |||

Db 361 CAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAG 412

# RESULT 7

BF108794

LOCUS BF108794 427 bp mRNA linear EST 20-OCT-2000  
 DEFINITION 7152g03.x1 Soares\_NSF\_F8\_9W\_OT\_PA\_P\_S1 Homo sapiens cDNA clone  
 IMAGE:3525292 3' similar to SW:NTAK\_HUMAN O14511 NTAK PROTEIN  
 ;contains MSR1.t1 MSR1 repetitive element ;, mRNA sequence.

ACCESSION BF108794

VERSION BF108794.1 GI:10938484

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 427)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

This clone is available royalty-free through LLNL ; contact the  
 IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.

Seq primer: -40UP from Gibco

High quality sequence stop: 396.

## FEATURES

source

Location/Qualifiers

1. .427

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3525292"

/lab\_host="DH10B"

/clone\_lib="Soares\_NSF\_F8\_9W\_OT\_PA\_P\_S1"

/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with  
 a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;  
 Equal amounts of plasmid DNA from five normalized  
 libraries were mixed, and ss circles were made in vitro.  
 Following HAP purification, this DNA was used as tracer in  
 a subtractive hybridization reaction. The driver was  
 PCR-amplified cDNAs from pools of 5,000 clones made from  
 the same 5 libraries. The pools consisted of the following  
 libraries and cloneIDs: Soares NbHSF pool 1:  
 309384-310919, 323208-325895 Soares Nb2HP pool 1:  
 145032-147335, 147720-148103, 148872-149255, 15002 -  
 150407, 151176-152327 Soares Nb2HF8-9W pool 1:  
 758280-760583, 772104-774407 Soares NbHPA pool 1:  
 304776-306311, 320136-322823, 326280-326663 Soares NbHOT  
 pool 1: 723720-726407, 739080-740999 Subtraction by Bento  
 Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 40.6%; Score 363.8; DB 10; Length 427;  
 Best Local Similarity 91.3%; Pred. No. 5.4e-64;  
 Matches 386; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
      || || | || | | | | | | | | | | | | | |
Db      5 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 64

Qy      423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
      |||||
Db      65 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 124

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
      |||||
Db      125 ATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 184

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 602
      |||||
Db      185 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 244

Qy      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
      |||||
Db      245 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 304

Qy      663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
      |||||
Db      305 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 364

Qy      723 CACCCTGTCATCCTGGTTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTATTG 782
      |||||
Db      365 CACCCTGTCATCCTGGTTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTATTG 424

Qy      783 CGT 785
      |||
Db      425 CGT 427
  
```

# RESULT 8

BX529505

LOCUS BX529505 488 bp mRNA linear EST 27-JUN-2003

DEFINITION BX529505 NCI CGAP Mam3 Mus musculus cDNA clone IMAGp998N017639 ;  
 IMAGE:3153984, mRNA sequence.

ACCESSION BX529505

VERSION BX529505.1 GI:32297863

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 488)

AUTHORS Heil,O., Ebert,L., Neubert,P., Peters,M., Radelof,U., Schneider,D.  
 and Korn,B.

TITLE Mouse UnigeneSet - RZPD2

JOURNAL Unpublished (2003)

COMMENT Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany



Db 301 ||||| 301 GCCAAGTCCTACTGTGTGAATGGAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTC 360  
 QY 829 TCCTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873  
 Db 361 ||||| 361 TCCTGCAAATGTCCAAACGGATTCTTCGGACAGAGATGTTTGGAG 405

RESULT 9  
 BI410828/c  
 LOCUS BI410828 949 bp mRNA linear EST 14-AUG-2001  
 DEFINITION 602963734F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5119065 5',  
 mRNA sequence.  
 ACCESSION BI410828  
 VERSION BI410828.1 GI:15171751  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 949)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM11290 row: d column: 10  
 High quality sequence start: 28  
 High quality sequence stop: 919.  
 FEATURES  
 source Location/Qualifiers  
 1. .949  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="Czech II"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5119065"  
 /tissue\_type="pooled lung tumors"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_Lu33"  
 /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: NotI; Site\_2: EcoRI; 1st  
 strand cDNA was prepared from mRNA obtained from pooled  
 lung tumors with a Not I - oligo(dT) primer [5'  
 TGTTACCAATCTGAAGTGGGAGCGGCCCTCTGTTTTTTTTTTTTTTT 3'].  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not  
 I and Eco RI sites of the modified pT7T3 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo. "



# ORIGIN

Query Match 28.4%; Score 255; DB 12; Length 949;  
 Best Local Similarity 80.0%; Pred. No. 1.3e-41;  
 Matches 388; Conservative 0; Mismatches 85; Indels 12; Gaps 7;

Qy	397	GTGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGC	456
Db	947	GTGGGCCAGATCCTGGGCACTG-CTGCGCCACCCGCCCAA-CTGAAGAAGATGAAGA-C	891
Qy	457	CAGACGGGACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCC	516
Db	890	CAAACCAGAAGAGTCGGTGAGAACAGTTCGCTCAAGTGTGAGGCACGGCCGGGGAAACCC	831
Qy	517	CAGCCTTCCTACC-----GTTGGTTCAAGGATGGC-AAGGAGCTCAACCGCAGCCGAGAC	570
Db	830	CCCCACCCCTCCCTATCGCTGGTTTCAAGGATGGCAAAGGAAGTCAACCGGAGTCGTGAT	771
Qy	571	ATTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAG	630
Db	770	ATTCGCATCAAGTATGCCAATGGCAGAAAGAACTCACGGCTACAGTTCAACAAAAGTGAG	711
Qy	631	GT--GGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCG	688
Db	710	GTTGGAGGATTGCCGGGGAGTACGTCTGTGAGGCCGAGAACATCCTTGGGGAAGGACACCG	651
Qy	689	TCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCATCCTGGTCGGGGCACG	748
Db	650	T-GAGGGCCGACTCCATGTCAACAGCGTGAGCACCCTCTGTCATCCTGGTCGGGGACATG	592
Qy	749	CCCGGAAGTGCAACGAGACAGCCAAGTCCATTGCGTCAATGGAGGCGTCTGCTACTACA	808
Db	591	CCCGGAAGTGCAATGAGACCGCCAAGTCTACTGTGTGAATGGAGGCGTGTGCTACTACA	532
Qy	809	TCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGTC	868
Db	531	TCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAACGGATTCTTCGGACAGAGATGTT	472
Qy	869	AGCAG	873
Db	471	TGGAG	467

RESULT 10  
 BE983573  
 LOCUS BE983573 333 bp mRNA linear EST 29-APR-2002  
 DEFINITION UI-M-CG0p-bgi-c-07-0-UI.s1 NIH\_BMAP\_Ret4\_S2 Mus musculus cDNA clone  
 UI-M-CG0p-bgi-c-07-0-UI 3', mRNA sequence.  
 ACCESSION BE983573  
 VERSION BE983573.1 GI:10654893  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 333)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: mEST@mail.nih.gov  
Oligo-dT track not found, Not I site shown in beginning of sequence is likely internal to the message. cDNA Library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements The tissue for this library was contributed by Dr. Xin-Yuan Fu, Yale University School of Medicine The following repetitive elements were found in this cDNA sequence: 15-105,  
>GC\_rich#Low\_complexity  
Seq primer: M13 Forward  
POLYA=No.

FEATURES

source Location/Qualifiers

1. .333

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UI-M-CG0p-bgi-c-07-0-UI"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="NIH\_BMAP\_Ret4\_S2"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; The NIH\_BMAP\_Ret4\_S2 library is a subtracted library, ultimately derived from mouse retina tissue libraries at various stages of development. For a detailed description of the library from which this clone was derived, please visit our web site at brainest.eng.uiowa.edu. The tissue for this library was contributed by Dr. Xin-Yuan Fu, Yale University School of Medicine

TAG\_SEQ=None found"

#### ORIGIN

Query Match 26.1%; Score 234.4; DB 10; Length 333;  
Best Local Similarity 95.6%; Pred. No. 1.3e-37;  
Matches 241; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCTCGCTCGCCTGC 60  
|||||

Db 82 ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 141

QY 61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120  
|||||

Db 142 TACTCGCCCAGCCTCAAGTCGGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 201

QY 121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180  
 |||||  
 Db 202 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCAGCTCTAACAGCACCCGAGAGCCG 261  
 QY 181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240  
 |||||  
 Db 262 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 321  
 QY 241 GGGCTGCAGCGC 252  
 |||||  
 Db 322 GGGCTGCAGCGC 333

RESULT 11

AA772412

LOCUS AA772412 297 bp mRNA linear EST 31-DEC-1998

DEFINITION ai44e12.s1 Soares\_parathyroid\_tumor\_NbHPA Homo sapiens cDNA clone  
 1359886 3' similar to TR:P43328 P43328 NEU DIFFERENTIATION FACTOR  
 NDF04 ;, mRNA sequence.

ACCESSION AA772412

VERSION AA772412.1 GI:2824195

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 297)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:

[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

Possible reversed clone: similarity on wrong strand

Possible reversed clone: polyT not found

Insert Length: 667 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 267.

FEATURES

source

Location/Qualifiers

1. .297

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="1359886"

/tissue\_type="parathyroid tumor"

/dev\_stage="adult"

/lab\_host="DH10B (ampicillin resistant)"

/clone\_lib="Soares\_parathyroid\_tumor\_NbHPA"

[note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCACCAATTTTTTTTTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid adenomas was kindly provided by Dr. Stephen Marx, National Institute of Diabetes and Digestive and Kidney Diseases, NIH."

ORIGIN

[illegible]

RESULT 12

BI651936  
LOCUS BI651936 795 bp mRNA linear EST 12-SEP-2001  
DEFINITION 603298677F1 NCI\_CGAP\_Mam3 Mus musculus cDNA clone IMAGE:5339251 5',  
mRNA sequence.  
ACCESSION BI651936  
VERSION BI651936.1 GI:15566172  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 795)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)



LOCUS BE648780 259 bp mRNA linear EST 06-SEP-2000  
 DEFINITION UI-M-BH2.2-aop-b-12-0-UI.r1 NIH\_BMAP\_M\_S3.2 Mus musculus cDNA clone  
 UI-M-BH2.2-aop-b-12-0-UI 5', mRNA sequence.  
 ACCESSION BE648780  
 VERSION BE648780.1 GI:9974601  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 259)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8889548  
 COMMENT Contact: Chin, H  
 National Institute of Mental Health  
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
 20892-9643, USA  
 Tel: 301 443 1706  
 Fax: 301 443 9890  
 Email: mEST@mail.nih.gov  
 cDNA Library Preparation: M.B. Soares Lab Clone distribution:  
 Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It  
 should be noted that Bento Soares is generating a small number of  
 additional specialized non-redundant arrays of BMAP cDNAs whose  
 availability will be considered under appropriate and limited  
 collaborative arrangements  
 Seq primer: M13 Reverse.  
 FEATURES  
 source Location/Qualifiers  
 1. .259  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UI-M-BH2.2-aop-b-12-0-UI"  
 /dev\_stage="27-32 days"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NIH\_BMAP\_M\_S3.2"  
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
 polylinker; Site\_1: Not I; Site\_2: Eco RI; The  
 NIH\_BMAP\_M\_S3.2 library is a subtracted library of a  
 series, ultimately derived from a mixture of individually  
 tagged normalized libraries from ten regions of the mouse  
 brain (cerebellum, brain stems, olfactory bulbs,  
 hypothalamus, cortex, amygdala, basal ganglia, pineal  
 gland, striatum, hippocampus) after a series of  
 subtractions to reduce the representation of cDNAs from  
 which ESTs had already been generated. The following  
 serially subtracted libraries were generated in this  
 process: NIH\_BMAP\_M\_S3.2, NIH\_BMAP\_M\_S2, NIH\_BMAP\_M\_S1.  
 The subtracted library (NIH\_BMAP\_M\_S3.2) was constructed  
 as follows: PCRamplified cDNA inserts from NIH\_BMAP\_M\_S2  
 clones from which 3' ESTs had been derived was used as a  
 driver in a hybridization with the NIH\_BMAP\_M\_S2 library

in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (LifeTechnologies) to generate the NIH\_BMAP\_M\_S3.2 library. This procedure has been previously described (Bonardo, Lennon and Soares, Genome Research 6:791-806, 1996)"

# ORIGIN

```

Query Match          21.8%;  Score 195.2;  DB 10;  Length 259;
Best Local Similarity 90.1%;  Pred. No. 1.2e-29;
Matches 209;  Conservative 0;  Mismatches 23;  Indels 0;  Gaps 0;

Qy      663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
      | | ||||| ||||| ||||| ||||| || | ||||| ||||| |||||
Db      1 CCACAACATCCTTGGGAAAGACACCGTGAGGGGCGGACTCCATGTCAAGAGCGTGAGCAC 60

Qy      723 CACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
      ||| ||||| ||||| ||||| || ||||| ||||| ||||| ||||| ||
Db      61 CACTCTGTCATCCTGGTCGGGACATGCCCGGAAGTGCAATGAGACCGCCAAGTCCTACTG 120

Qy      783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      121 TGTGAATGGGGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 180

Qy      843 TGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGGTCAACTTCTCC 894
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      181 TGTGGGATACACCGGTGACAAGTGTGTCAGCAGTTCGCAATGGTCAACTTCTCC 232

```

# RESULT 14

AA968077

```

LOCUS      AA968077          327 bp    mRNA    linear    EST 19-MAY-1998
DEFINITION uh09h01.r1 Soares mouse hypothalamus NMHy Mus musculus cDNA clone
IMAGE:1617457 5' similar to TR:035073 035073 NTAK ALPHA2-1P ;, mRNA
sequence.
ACCESSION  AA968077
VERSION    AA968077.1  GI:3141970
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 327)
AUTHORS    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE      The WashU-HHMI Mouse EST Project
JOURNAL    Unpublished (1996)
COMMENT    Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800

```

Fax: 314 286 1810  
 Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LLNL ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:956757  
 Seq primer: -28ml3 rev2 ET from Amersham  
 High quality sequence stop: 318.

FEATURES  
 source Location/Qualifiers  
 1. .327  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:1617457"  
 /tissue\_type="hypothalamus"  
 /lab\_host="DH10B"  
 /clone\_lib="Soares mouse hypothalamus NMHy"  
 /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5'  
 TGTTACCAATCTGAAGTGGGAGCGGCCCAAGGTTTTTTTTTTTTTTTTTTTTT  
 T 3']; double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT7T3 vector. RNA  
 provided by Dr. Wolfgang Liedtke. Library went through  
 two rounds of normalization, and was constructed by Bento  
 Soares and M.Fatima Bonaldo."

#### ORIGIN

Query Match 21.8%; Score 195.2; DB 9; Length 327;  
 Best Local Similarity 90.2%; Pred. No. 1.4e-29;  
 Matches 231; Conservative 0; Mismatches 23; Indels 2; Gaps 2;

```

Qy      641 CTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGC 700
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      2   CCGGGGAGTACGTCTGTGAGGCCCAGAACATCCTTGGAAGGACACCGTGAGGGA-CGAC 60

Qy      701 TTTACGTCAACAGCGTGAGCACCACCCTGTTCATCCTGGTCGGGGCACGCCCGGAAGTGCA 760
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61   TCCATGTCAACAGCGTGAG-ACCACTCTGTTCATCCTGGTCGGGACATGCCCGGAAGTGCA 119

Qy      761 ACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCA 820
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      120 ATGAGACCGCCAAGTCCTACTGTGTGAATGGAGGCGTGTGCTACTACATCGAGGGCATCA 179

Qy      821 ACCAGCTCTCCTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAA 880
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      180 ACCAGCTCTCCTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAA 239

Qy      881 TGGTCAACTTCTCCTA 896
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      240 TGGTCAACATCTCCAA 255
  
```

#### RESULT 15

BX089049/c

LOCUS BX089049 362 bp mRNA linear EST 23-JAN-2003

DEFINITION BX089049 Soares\_parathyroid\_tumor\_NbHPA Homo sapiens cDNA clone



IMAGp998M133119 ; IMAGE:1240116, mRNA sequence.  
 ACCESSION BX089049  
 VERSION BX089049.1 GI:27825909  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 362)  
 AUTHORS Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,  
 Radelof,U., Schneider,D. and Korn,B.  
 TITLE Human UnigeneSet - RZPD3  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Ina Rolfs  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
 RZPD; IMAGp998M133119.  
 RZPDLIB; I.M.A.G.E. cDNA Clone Collection;  
 Human UnigeneSet - RZPD3 (RZPDLIB No.972)  
 http://www.rzpd.de/CloneCards/cgi-  
 bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Heubnerweg 6, D-14059 Berlin, Germany  
 Tel: +49 30 32639 101  
 Fax: +49 30 32639 111  
 www.rzpd.de  
 This clone is available royalty-free from RZPD;  
 contact RZPD (clone@rzpd.de) for further information. Seq primer:  
 M13r, Primer sequence: TTTACACAGGAAACAGCTATGAC.  
 FEATURES  
 source Location/Qualifiers  
 1. .362  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGp998M133119 ; IMAGE:1240116"  
 /tissue\_type="parathyroid tumor"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares\_parathyroid\_tumor\_NbHPA"  
 /note="Organ: parathyroid gland; Vector: pT7T3D  
 (Pharmacia) with a modified polylinker; Site\_1: Not I;  
 Site\_2: Eco RI; 1st strand cDNA was primed with a Not I -  
 oligo(dT) primer  
 [5'-TGTTACCAATCTGAAGTGGGAGCGCGCCGACCAATTTTTTTTTTTTTTTTTT  
 TTTT-3'], double-stranded cDNA was size selected, ligated  
 to Eco RI adapters (Pharmacia), digested with Not I and  
 cloned into the Not I and Eco RI sites of a modified pT7T3  
 vector (Pharmacia). Library went through one round of  
 normalization to a Cot = 5. Library constructed by Bento  
 Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid  
 adenomas was kindly provided by Dr. Stephen Marx, National  
 Institute of Diabetes and Digestive and Kidney Diseases,  
 NIH."  
 ORIGIN  
 Query Match 20.4%; Score 182.6; DB 13; Length 362;  
 Best Local Similarity 85.3%; Pred. No. 5.3e-27;

Matches 203; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy	656	GCGAGGCCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCG	715
Db	362	GCGAGGCCCGAGAACATCCTGGGGAAGGACACCGTCCGGGNCCGGCTTTACGTCAACAGCG	303
Qy	716	TGAGCACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGT	775
Db	302	TGAGCACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGT	243
Qy	776	CCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCA	835
Db	242	CCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCA	183
Qy	836	AGTGTCTCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGGTCAACTTCTC	893
Db	182	AGGCACATGGGCTGCACTGCTTAGAAGTTGGTACCCAGAGCCACCACTTCCCCATCTC	125

Search completed: August 15, 2004, 09:42:30  
Job time : 2850.01 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 02:48:13 ; Search time 3828.5 Seconds  
(without alignments)  
10155.083 Million cell updates/sec

Title: US-09-864-675-3  
Perfect score: 897  
Sequence: 1 atgaggcgcgacccggcccc.....caatggtcaacttctcctaa 897

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*

28: em\_un:\*  
 29: em\_vi:\*  
 30: em\_htg\_hum:\*  
 31: em\_htg\_inv:\*  
 32: em\_htg\_other:\*  
 33: em\_htg\_mus:\*  
 34: em\_htg\_pln:\*  
 35: em\_htg\_rod:\*  
 36: em\_htg\_mam:\*  
 37: em\_htg\_vrt:\*  
 38: em\_sy:\*  
 39: em\_htgo\_hum:\*  
 40: em\_htgo\_mus:\*  
 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result		Query					Description
No.	Score	Match	Length	DB	ID		
	1	849	94.6	3020	9	AB005060	AB005060 Homo sapi
	2	847.4	94.5	2613	9	AK124504	AK124504 Homo sapi
	3	835.4	93.1	1884	6	AR098145	AR098145 Sequence
	4	835.4	93.1	1884	6	AR116617	AR116617 Sequence
	5	784	87.4	3441	6	AR072052	AR072052 Sequence
	6	738.6	82.3	993	6	AR072053	AR072053 Sequence
	7	737	82.2	2947	10	D89995	D89995 Rattus sp.
	8	737	82.2	3076	6	E16456	E16456 Rat mRNA fo
	9	737	82.2	3077	10	D89996	D89996 Rattus sp.
	10	427.8	47.7	1476	6	AR098146	AR098146 Sequence
	11	427.8	47.7	1476	6	AR116618	AR116618 Sequence
	12	427.8	47.7	2268	6	AR098155	AR098155 Sequence
	13	427.8	47.7	2268	6	AR116627	AR116627 Sequence
	14	425	47.4	2188	10	AB001576	AB001576 Rattus sp
	15	424.8	47.4	118504	9	AC094080	AC094080 Homo sapi
c	16	424.8	47.4	152838	2	AC011589	AC011589 Homo sapi
	17	424.8	47.4	170797	9	AC011379	AC011379 Homo sapi
	18	424.8	47.4	210675	2	AC026272	AC026272 Homo sapi
	19	424	47.3	1054	6	AX406616	AX406616 Sequence
	20	424	47.3	1054	9	HS2NRG01	AF119151 Homo sapi
	21	405.4	45.2	1607	6	AR098144	AR098144 Sequence
	22	405.4	45.2	1607	6	AR116616	AR116616 Sequence
	23	405.4	45.2	2467	6	AR098143	AR098143 Sequence
	24	405.4	45.2	2467	6	AR116615	AR116615 Sequence
	25	387.2	43.2	139074	2	AC131191	AC131191 Mus muscu
	26	384	42.8	253462	2	AC096477	AC096477 Rattus no
	27	216.2	24.1	1207	6	AR072054	AR072054 Sequence
	28	173	19.3	419	6	AX406617	AX406617 Sequence
	29	173	19.3	419	9	HS2NRG02	AF119152 Homo sapi
	30	173	19.3	120236	9	AC008523	AC008523 Homo sapi
c	31	173	19.3	189049	9	AC008667	AC008667 Homo sapi
	32	142.6	15.9	85703	2	AC020830	AC020830 Mus muscu
	33	142.6	15.9	190228	10	AC127350	AC127350 Mus muscu

	34	139.4	15.5	226038	2	AC106592	AC106592 Rattus no
	35	139.4	15.5	273080	2	AC098540	AC098540 Rattus no
c	36	139.4	15.5	302176	2	AC096479	AC096479 Rattus no
	37	130.2	14.5	163	10	AY227026	AY227026 Mus muscu
	38	124.6	13.9	493	6	AX406618	AX406618 Sequence
	39	124.6	13.9	493	9	HS2NRG03	AF119153 Homo sapi
	40	124	13.8	350	6	AX406619	AX406619 Sequence
	41	124	13.8	350	9	HS2NRG04	AF119154 Homo sapi
c	42	109.6	12.2	85703	2	AC020830	AC020830 Mus muscu
c	43	108.4	12.1	205280	2	BX323592	BX323592 Danio rer
	44	108.4	12.1	207840	5	BX005008	BX005008 Zebrafish
	45	97	10.8	172	10	D89997	D89997 Rattus sp.

# ALIGNMENTS

## RESULT 1

AB005060

LOCUS AB005060 3020 bp mRNA linear PRI 14-NOV-1997

DEFINITION Homo sapiens mRNA for NTAK, complete cds.

ACCESSION AB005060

VERSION AB005060.1 GI:2626738

KEYWORDS NTAK.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (sites)

AUTHORS Higashiyama,S., Horikawa,M., Yamada,K., Ichino,N., Nakano,N., Nakagawa,T., Miyagawa,J., Matsushita,N., Nagatsu,T., Taniguchi,N. and Ishiguro,H.

TITLE A novel brain-derived member of the epidermal growth factor family that interacts with ErbB3 and ErbB4

JOURNAL J. Biochem. 122 (3), 675-680 (1997)

MEDLINE 98006324

PUBMED 9348101

REFERENCE 2 (bases 1 to 3020)

AUTHORS Ishiguro,H.

TITLE Direct Submission

JOURNAL Submitted (24-JUN-1997) Hiroshi Ishiguro, Fujita Health University, ICMS; 1-98, kutsukake-cho, Toyoake, Aichi 470-11, Japan (E-mail:hishi@fujita-hu.ac.jp, Tel:0562-93-9393, Fax:0562-93-8831)

FEATURES Location/Qualifiers

source 1..3020

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/cell\_line="SK-N-SH"

/cell\_type="neuroblastoma"

CDS 226..2778

/codon\_start=1

/product="NTAK"

/protein\_id="BAA23417.1"

/db\_xref="GI:2626739"

/translation="MRQVCCSALPPPPLEKGRCSYSDSSSSSSERSSSSSSSSSSESG  
SSSRSSNNSSISRPAAPPEPRPQQPQPRSPAARRAAARSRAAAAGGMRRDPAPGFS"

MLLEGVSLACYSPSLKSVQDQAYKAPVVVEGKVQGLVPAGGSSSNSTREPPASGRVAL  
VKVLDKWPLRSGGLQREQVISVGSCVPLERNQRYIFFLEPTEQPLVFKTAFAPLDTNG  
KNLKKEVGKILCTDCATRPKLKMKMSQTGQVGEKQSLKCEAAAGNPQPSYRWFKDGKE  
LNRSRDIRIKYGNRKNRSLQFNKVKVEDAGEYVCEAENILGKDTVGRGLVNSVSTT  
LSSWSGHARKCNETAKSYCVNGGVCIYIEGINQLSCKCPNGFFGQRCLEKLPLRLYMP  
DEKQKAEELYQKRVLTTITGICVALLVVGIVCVVAYCKTKKQKQMHNLHRQNMCPAHQ  
NRSLANGPSHPRLDPPEIQMADYISKNPATDHVIRRETETTFSGSHSCSPSHHCSTA  
TPTSSHRHESHTWSLSESLTSDSQSGIMLSSVGTSKCNTPACVEARARRAAAYNLE  
ERRRATAPPYHDSVDSLDRDSPHSERYVSALTTPARLSPVDFHYSLATQVPTFEITSPN  
SAHAVSLPPAAPISYRLAEQQPLLRHPAPPGPGPGPGPGPGPGADMQRSYDSYYYPAA  
GPGPRRGTCALGGSGLSLPASPFRIPEDEYETTQECAPPPPPRPRARGASRRTSAGP  
RRWRRSRLNGLAAQARAARDSLSLSGSGGGSASASDDDDADDADGALAAESTPFLGL  
RGAHDALRSDSPPLCPAADSRTYYSLDSHSTRASSRHSRGPPPPRAKQDSAPL"

polyA\_site

3020

/note="39 A nucleotides"

ORIGIN

Query Match 94.6%; Score 849; DB 9; Length 3020;  
Best Local Similarity 98.3%; Pred. No. 5.5e-152;  
Matches 858; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC	60
Db	502	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC	561
Qy	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	562	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	621
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	622	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	681
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	682	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	741
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	742	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	801
Qy	301	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	802	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	861
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	862	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	921
Qy	421	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	922	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	981
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	982	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	1041

Qy 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 600  
 |||  
 Db 1042 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 1101  
 Qy 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660  
 |||  
 Db 1102 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 1161  
 Qy 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720  
 |||  
 Db 1162 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 1221  
 Qy 721 ACCACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780  
 |||  
 Db 1222 ACCACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 1281  
 Qy 781 TCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840  
 |||  
 Db 1282 TCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 1341  
 Qy 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873  
 || ||| | ||| ||| |||  
 Db 1342 CCAAATGGATTCTTCGGACAGAGATGTTTGGAG 1374

## RESULT 2

AK124504

LOCUS AK124504 2613 bp mRNA linear PRI 09-SEP-2003

DEFINITION Homo sapiens cDNA FLJ42513 fis, clone BRACE2046295, highly similar to NTAK PROTEIN.

ACCESSION AK124504

VERSION AK124504.1 GI:34530302

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1

AUTHORS Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2613)

AUTHORS Isogai,T. and Yamamoto,J.

TITLE Direct Submission

JOURNAL Submitted (15-JUL-2003) Takao Isogai, FLJ Project(HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan

Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES                      Location/Qualifiers  
    source                    1. .2613  
                              /organism="Homo sapiens"  
                              /mol\_type="mRNA"  
                              /db\_xref="taxon:9606"  
                              /clone="BRACE2046295"  
                              /tissue\_type="cerebellum"  
                              /clone\_lib="BRACE2"  
                              /note="cloning vector: pME18SFL3"

ORIGIN

Query Match                      94.5%;    Score 847.4;    DB 9;    Length 2613;  
Best Local Similarity    98.2%;    Pred. No. 1.1e-151;  
Matches 857;    Conservative    0;    Mismatches    16;    Indels        0;    Gaps        0;

```
Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60
          |||
Db    436 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTTGCTCGCCTGC 495

Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
          |||
Db    496 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 555

Qy     121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
          |||
Db    556 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 615

Qy     181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
          |||
Db    616 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 675

Qy     241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
          |||
Db    676 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 735

Qy     301 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
          |||
Db    736 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 795

Qy     361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
          |||
Db    796 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 855

Qy     421 TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
          |||
Db    856 TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 915

Qy     481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
          |||
Db    916 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 975

Qy     541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 600
          |||
```



Db 976 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 1035

Qy 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660  
 |||

Db 1036 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 1095

Qy 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720  
 |||

Db 1096 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 1155

Qy 721 ACCACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780  
 |||

Db 1156 ACCACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 1215

Qy 781 TCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840  
 |||

Db 1216 TCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAATGT 1275

Qy 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873  
 || ||| | || | || | ||

Db 1276 CCAAATGGATTCTTCGGACAGAGATGTTTGGAG 1308

# RESULT 3

AR098145

LOCUS AR098145 1884 bp DNA linear PAT 14-FEB-2001

DEFINITION Sequence 5 from patent US 6074841.

ACCESSION AR098145

VERSION AR098145.1 GI:12807402

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1884)

AUTHORS Gearing, D.P. and Busfield, S.J.

TITLE Don-1 gene and polypeptides and uses therefor

JOURNAL Patent: US 6074841-A 5 13-JUN-2000;

FEATURES Location/Qualifiers

source 1. .1884

/organism="unknown"

/mol\_type="unassigned DNA"

## ORIGIN

Query Match 93.1%; Score 835.4; DB 6; Length 1884;

Best Local Similarity 98.1%; Pred. No. 2.2e-149;

Matches 856; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

Qy 1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60  
 |||

Db 218 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 277

Qy 61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120  
 |||

Db 278 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 337

Qy 121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180  
 |||

Db	338	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	397
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	398	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	457
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	458	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	517
Qy	301	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	518	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	577
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCAC'TGAC	420
Db	578	CCCCT-GATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCAC'TGAC	636
Qy	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	637	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	696
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	697	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	756
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	757	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG	816
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	817	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	876
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	877	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	936
Qy	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	937	ACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	996
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAAGTGT	840
Db	997	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	1056
Qy	841	CCTGTGGGATACACCGGGGACAGGTGTCAGCAG	873
Db	1057	CCAAATGGATTCTTCGGACAGAGATGTTTGGAG	1089

#### RESULT 4

AR116617

LOCUS AR116617 1884 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 5 from patent US 6133423.

ACCESSION AR116617

VERSION AR116617.1 GI:14096939



Qy 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660  
 |||  
 Db 817 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 876  
 Qy 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720  
 |||  
 Db 877 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 936  
 Qy 721 ACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780  
 |||  
 Db 937 ACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 996  
 Qy 781 TCGGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840  
 |||  
 Db 997 TCGGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAATGT 1056  
 Qy 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873  
 || ||| | ||| | || ||| | ||  
 Db 1057 CCAAATGGATTCTTCGGACAGAGATGTTTGGAG 1089

# RESULT 5

AR072052

LOCUS AR072052 3441 bp DNA linear PAT 18-FEB-2000

DEFINITION Sequence 1 from patent US 5912326.

ACCESSION AR072052

VERSION AR072052.1 GI:7222940

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 3441)

AUTHORS Chang,H.

TITLE Cerebellum-derived growth factors

JOURNAL Patent: US 5912326-A 1 15-JUN-1999;

FEATURES Location/Qualifiers

source 1. .3441

/organism="unknown"

/mol\_type="unassigned DNA"

## ORIGIN

Query Match 87.4%; Score 784; DB 6; Length 3441;

Best Local Similarity 92.2%; Pred. No. 1.4e-139;

Matches 826; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60  
 |||  
 Db 180 ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 239  
 Qy 61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120  
 |||  
 Db 240 TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 299  
 Qy 121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180  
 |||  
 Db 300 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT 359



SOURCE Unknown.  
 ORGANISM Unknown.  
 Unclassified.  
 REFERENCE 1 (bases 1 to 993)  
 AUTHORS Chang,H.  
 TITLE Cerebellum-derived growth factors  
 JOURNAL Patent: US 5912326-A 3 15-JUN-1999;  
 FEATURES Location/Qualifiers  
     source 1. .993  
             /organism="unknown"  
             /mol\_type="unassigned DNA"

ORIGIN

Query Match 82.3%; Score 738.6; DB 6; Length 993;  
 Best Local Similarity 90.4%; Pred. No. 6.8e-131;  
 Matches 789; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC	60
Db	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC	60
Qy	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	61	TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	121	GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT	180
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	181	CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG	300
Qy	301	CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	301	CGCTACATCTTTTTCTGGAGCCACCGAGCAGCCCTTAGTTTTTAAGACAGCCTTTGCC	360
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	361	CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Qy	421	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	421	TGCGCAACCCGGCCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG	480
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCTACCGTTGGTTCAAG	540
Db	481	CAGTCGCTCAAGTGTGAGGCGGCGGGGAACCCCCAGCCCTCCTATCGATGGTTCAAG	540
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAG	600
Db	541	GACGGCAAGGAGCTCAACCGGAGTCGTGACATTGCGATCAAGTATGGCAACGGCAGAAAG	600



```
/codon_start=1
/product="NTAK alpha1"
/protein_id="BAA23344.1"
/db_xref="GI:2605630"
/translation="MRQVCCSALPPPLEKARCSSYSYSDSSSSSSSSNNSSSSTSSRSS
SRSSSRSSRGSTTTTSSSENSGNSGSI FRPAAPPEPRPQPQPRSPAARRAAARSR
AAAAGGMRRDPAPGSSMLLFGVSLACYSPLKSVQDQAYKAPVVVEGKVQGLAPAGGS
SSNSTREPPASGRVALVKVLDKWPLRSGGLQREQVISVGSCAPLERNQRYIFFLEPTE
QPLVFKTAFAFVDPNGKNIKKEVGKILCTDCATRPKLKKMKSQTGEVGEKQSLKCEAA
AGNPQPSYRWFKDGKELNRSRDIRIKYGNRKN SRLQFNKVKVEDAGEYVCEAENILG
KDTVGRGLHVNVSSTTLSSWSGHARKCNETAKSYCVNGGV CYIIEGINQLSCKCPNGF
FGQRCLEKLPLRLYMPDPKQKHLGFELKEAEELYQKRVLTTITGICVALLVGVICVVA
YCKTKKQRRQMHHHLRQNMCPAHQNRSLANGP SHPRLDPEEIQMADYISKNPATD HV
IRREAETTFSGSHSCSPSHHCSTATPTSSHRHESHTWSLERSESLTSDSQSGIMLSSV
GTSKCN SPACVEARARRAAAYSQEERRRAAMP PYHDSIDSLRDSPHSERYVSALTTPA
RLSPVDFHYSLATQVPTFEITSPNSAHAVSLPPAAPISYRLAEQQPLL RHPAPPGPGP
GPGADMQRSYDSYYP AAGPGPRRGACALGGS LGSLPASPFRI PEDDEYETTQECAPP
PPPRPRTRGASRRTSAGPRRWRRSRLNGLAAQRARAARDSLSLSSGSGCGSASASDDD
ADDADGALAAESTPFLGLRAAHDALRSDSPPLCPAADSRTYYS LDSHSTRASSRHSRG
PPTRAKQDSGPL"
```

ORIGIN

Query Match 82.2%; Score 737; DB 10; Length 2947;  
Best Local Similarity 90.3%; Pred. No. 1.3e-130;  
Matches 788; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```
Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60
      |||
Db    403 ATGAGGCGCGACCCGGCCCCCGGCTCCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 462

Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
      |||
Db    463 TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 522

Qy     121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
      |||
Db    523 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT 582

Qy     181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
      |||
Db    583 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 642

Qy     241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
      |||
Db    643 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG 702

Qy     301 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
      |||
Db    703 CGCTACATCTTTTTCCTGGAGCCACCGAGCAGCCCTTAGTTTTTAAGACAGCCTTTGCC 762

Qy     361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
      ||
Db    763 CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 822

Qy     421 TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
      |||
Db    823 TGCGCAACCCGGCCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG 882
```



Qy 481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540  
 || ||||| ||||| || || || || ||||| ||||| || ||||| ||  
 Db 883 CAGTCGCTCAAGTGTGAGGCGGCGGCGGGGAACCCCCAGCCCTCCTATCGATGGTTCAAG 942  
 Qy 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 600  
 || ||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||  
 Db 943 GACGGCAAGGAGCTCAACCGGAGTCGTGACATTTCGCATCAAGTATGGCAACGGCAGAAAG 1002  
 Qy 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
 Db 1003 AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG 1062  
 Qy 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720  
 || ||||| ||||| ||||| || ||||| ||||| ||||| ||||| ||  
 Db 1063 GCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGTGAGC 1122  
 Qy 721 ACCACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
 Db 1123 ACCACTCTGTCTGCTCCTGGTCGGGGCAGCCCCGGAAGTGCAATGAGACAGCCAAGTCCTAC 1182  
 Qy 781 TCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840  
 || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||  
 Db 1183 TGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAATGT 1242  
 Qy 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873  
 || ||||| || || || || || || || ||  
 Db 1243 CCAAACGGATTCTTCGGACAGAGATGTTTGGAG 1275

# RESULT 8

E16456

LOCUS E16456 3076 bp DNA linear PAT 28-JUL-1999

DEFINITION Rat mRNA for neuregulin-like Transmembrane Activator for ErbB Kinases (NTAK).

ACCESSION E16456

VERSION E16456.1 GI:5711139

KEYWORDS JP 1998179166-A/1.

SOURCE Rattus sp.

ORGANISM Rattus sp.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 3076)

AUTHORS Higashiyama,S., Taniguchi,N., Ishiguro,K. and Nagatsu,T.

TITLE GENE ENCODING RECEPTOR TYPE TYROSINE-KINASE ERB B LIGAND AND ITS

JOURNAL Patent: JP 1998179166-A 1 07-JUL-1998;

HIGASHIYAMA SHIGEKI

COMMENT OS Rattus sp. (rat)

PN JP 1998179166-A/1

PD 07-JUL-1998

PF 25-DEC-1996 JP 1996356998

PI HIGASHIYAMA SHIGEKI, TANIGUCHI NAOYUKI, ISHIGURO KEIJI, PI NAGATSU TOSHIHARU

PC C12N15/09,C07K14/705,C07K16/28,C12N5/10,C12N15/02,C12P21/02,

PC C12P21/08,

PC C12Q1/68,G01N33/53,G01N33/566//A61K31/70,A61K38/46,A61K39/395,

```

PC      A61K48/00,
PC      C07H21/04, (C12N5/10,C12R1:91), (C12P21/02,C12R1:91); CC
strandedness: Double;
CC      topology: Linear;
FH      Key                      Location/Qualifiers
FH
FT      source                    1. .3076
FT                                  /organism='Rattus sp.'
FT                                  /cell_line='PC12'
FT      CDS                      232. .2814
FT                                  /product='NTAK protein'.

FEATURES
      source                    Location/Qualifiers
                                1. .3076
                                /organism="Rattus sp."
                                /mol_type="genomic DNA"
                                /db_xref="taxon:10118"

ORIGIN

Query Match      82.2%;   Score 737;   DB 6;   Length 3076;
Best Local Similarity  90.3%;   Pred. No. 1.3e-130;
Matches 788;   Conservative  0;   Mismatches  85;   Indels  0;   Gaps  0;

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCTCGCTCGCCTGC 60
      ||||||||||||||||||||||||| ||| ||||||||||||||||||| |||||||||
Db      556 ATGAGGCGCGACCCGGCCCCCGGCTCCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 615

Qy      61 TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
      ||||||||||||||||||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      616 TACTCGCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 675

Qy      121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
      ||||||||||||||| ||||| ||| || ||||| ||||| ||||| ||||| ||||| |||||
Db      676 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCAGCTCTAACAGCACCCGAGAGCCT 735

Qy      181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
      ||||||||||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      736 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 795

Qy      241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
      ||||||||||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      796 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG 855

Qy      301 CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
      ||||||||||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      856 CGCTACATCTTTTTCTGGAGCCACCGAGCAGCCCTTAGTTTTTAAGACAGCCTTTGCC 915

Qy      361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
      || ||||| | ||||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      916 CCGGTCGACCTTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 975

Qy      421 TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
      ||||| ||||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      976 TGCGCAACCCGGCCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG 1035

Qy      481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
      || ||||| ||||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1036 CAGTCGCTCAAGTGTGAGGCGGCGGGGAACCCCCAGCCCTCCTATCGATGGTTCAAG 1095

```

Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	1096	GACGGCAAGGAGCTCAACCGGAGTCGTGACATTTCGCATCAAGTATGGCAACGGCAGAAAG	1155
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	1156	AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG	1215
Qy	661	GCCGAGAACATCCTTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	1216	GCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGTGAGC	1275
Qy	721	ACCACCCTGTCATCCTGGTTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	1276	ACCACTCTGTCTGTCCTGGTTCGGGGCACGCCCGGAAGTGCAATGAGACAGCCAAGTCCTAC	1335
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT	840
Db	1336	TGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAAATGT	1395
Qy	841	CCTGTGGGATACACCGGGGACAGGTGTCAGCAG	873
Db	1396	CCAAACGGATTCTTCGGACAGAGATGTTTGGAG	1428

CDS

```
/db_xref="taxon:10118"  
/cell_line="PC12"  
/cell_type="pheochromocytoma"  
233. .2815  
/codon_start=1  
/product="NTAK alpha2"  
/protein_id="BAA23345.1"  
/db_xref="GI:2605632"  
/translation="MRQVCCSALPPPLEKARCSSYSYSDSSSSSSSSNNSSSTSSRSS  
SRSSSRSSRGSTTTTSSSENSGSNSGSI FRPAAPPEPRPQPQPQRSPAARRAAARSR  
AAAAGGMRRDPAPGSSMLLFVSLACYSPLKSVQDQAYKAPVVVEGKVQGLAPAGGS  
SSNSTREPPASGRVALVKVLDKWPLRSGGLQREQVISVGSCAPLERNQRYIFFLEPTE  
QPLVFKTAFAPVDPNGKNIKKEVGKILCTDCATRPKLKMKMSQTGEVGEKQSLKCEAA  
AGNPQPSYRWFKDGKELNRSRDIRIKYGNRKNRSLQFNKVKVEDAGEYVCEAENILG  
KDTVGRGLHVNSVSTTLSSWSGHARKCNETAKSYCVNGGVCIYIEGINQLSCKCPNGF  
FGQRCLEKLPLRLYMPDPKQKAEELYQKRVLTTITGICVALLVVGIVCVVAYCKTKQR  
RQMHHHLRQNMCPAHQNRSLANGPSHPRLDPEEIQMADYISKNPATDHVIRREAETT  
FSGSHSCSPSHHCSTATPTSSHRHESHTWSLSESLTSDSQQSGIMLSSVGTSKCNSP  
ACVEARARRAAAYSQEERRRAAMPYPYHDSIDSLRDSPHSERYVSALTTPARLSPVDFH  
YSLATQVPTFEITSPNSAHAVSLPPAAPISYRLAEQQPLL RHPAPPGP GPGADMQR  
SYDSYYYP AAGP GPRRGACALGGS LGSLPASPFRIPEDEYETTQECAPPPPPRPRTR  
GASRRTSAGPRRWRRSRLNGLAAQRARAARDSLSLSSGSGCGSASASDDDADDADGAL  
AAESTPFLGLRAAHDALRSDSPPLCPAADSRTYYSLDSHSTRASSRHSRGPPTRAKQD  
SGPL"
```

ORIGIN

Query Match 82.2%; Score 737; DB 10; Length 3077;  
Best Local Similarity 90.3%; Pred. No. 1.3e-130;  
Matches 788; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```
Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60  
      |||  
Db      557 ATGAGGCGCGACCCGGCCCCCGGCTCCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 616  
      |||  
Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120  
      |||  
Db      617 TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 676  
      |||  
Qy      121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180  
      |||  
Db      677 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT 736  
      |||  
Qy      181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240  
      |||  
Db      737 CCCGCCTCGGGTCGGGTGGCGTTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 796  
      |||  
Qy      241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300  
      |||  
Db      797 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG 856  
      |||  
Qy      301 CGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360  
      |||  
Db      857 CGCTACATCTTTTTCCTGGAGCCCACCGAGCAGCCCTTAGTCTTTAAGACAGCCTTTGCC 916  
      |||  
Qy      361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420  
      ||  
Db      917 CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 976  
      ||
```



QY 363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422  
 || || | || | | | | || | | | | ||  
 Db 98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 157

QY 423 CGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482  
 |||||  
 Db 158 AGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

QY 483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542  
 |||||  
 Db 218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

QY 543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 602  
 |||||  
 Db 278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 337

QY 603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662  
 |||||  
 Db 338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

QY 663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722  
 |||||  
 Db 398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

QY 723 CACCCTGTCATCCTGGTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTATTG 782  
 |||||  
 Db 458 CACCCTGTCATCCTGGTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTATTG 517

QY 783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842  
 |||||  
 Db 518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAATGTCC 577

QY 843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873  
 |||| | || | || || ||  
 Db 578 AAATGGATTCTTCGGACAGAGATGTTTGGAG 608

# RESULT 11

AR116618

LOCUS AR116618 1476 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 7 from patent US 6133423.

ACCESSION AR116618

VERSION AR116618.1 GI:14096940

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 1476)

AUTHORS Gearing,D.P. and Busfield,S.J.

TITLE Don-1 gene and polypeptides and uses therefor

JOURNAL Patent: US 6133423-A 7 17-OCT-2000;

FEATURES Location/Qualifiers

source 1..1476

/organism="unknown"

/mol\_type="unassigned DNA"

ORIGIN

Query Match 47.7%; Score 427.8; DB 6; Length 1476;  
 Best Local Similarity 89.8%; Pred. No. 1.3e-71;  
 Matches 459; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
          || || | || | | | | || | | | | | | |
Db      98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
          |||||
Db      158 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
          |||||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAGAA 602
          |||||
Db      278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAGAA 337

Qy      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
          |||||
Db      338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

Qy      663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
          |||||
Db      398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

Qy      723 CACCCTGTCATCCTGGTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTATTG 782
          |||||
Db      458 CACCCTGTCATCCTGGTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTATTG 517

Qy      783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842
          |||||
Db      518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAATGTCC 577

Qy      843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873
          |||| | || | || | ||
Db      578 AAATGGATTCTTCGGACAGAGATGTTTGGAG 608
  
```

# RESULT 12

AR098155

LOCUS AR098155 2268 bp DNA linear PAT 14-FEB-2001

DEFINITION Sequence 31 from patent US 6074841.

ACCESSION AR098155

VERSION AR098155.1 GI:12807412

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 2268)

AUTHORS Gearing,D.P. and Busfield,S.J.

TITLE Don-1 gene and polypeptides and uses therefor

JOURNAL Patent: US 6074841-A 31 13-JUN-2000;

FEATURES Location/Qualifiers

```

source
1. .2268
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match          47.7%;   Score 427.8;   DB 6;   Length 2268;
Best Local Similarity 89.8%;   Pred. No. 1.2e-71;
Matches 459;   Conservative 0;   Mismatches 52;   Indels 0;   Gaps 0,

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
      || || | || | | | | || | | | | | | | |
Db      98  CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
      |||||
Db      158 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
      |||||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 602
      |||||
Db      278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 337

Qy      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
      |||||
Db      338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

Qy      663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
      |||||
Db      398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

Qy      723 CACCCTGTCATCCTGGTCGGGGCACGCCCGBAAGTGCAACGAGACAGCCAAGTCCTATTG 782
      |||||
Db      458 CACCCTGTCATCCTGGTCGGGGCACGCCCGBAAGTGCAACGAGACAGCCAAGTCCTATTG 517

Qy      783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842
      |||||
Db      518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAATGTCC 577

Qy      843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873
      |||| | || | || | || |
Db      578 AAATGGATTCTTCGGACAGAGATGTTTGGAG 608

```





SOURCE Rattus sp.  
 ORGANISM Rattus sp.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (sites)  
 AUTHORS Higashiyama,S., Horikawa,M., Yamada,K., Ichino,N., Nakano,N.,  
 Nakagawa,T., Miyagawa,J., Matsushita,N., Nagatsu,T., Taniguchi,N.  
 and Ishiguro,H.  
 TITLE A novel brain-derived member of the epidermal growth factor family  
 that interacts with ErbB3 and ErbB4  
 JOURNAL J. Biochem. 122 (3), 675-680 (1997)  
 MEDLINE 98006324  
 PUBMED 9348101  
 REFERENCE 2 (bases 1 to 2188)  
 AUTHORS Ishiguro,H.  
 TITLE Direct Submission  
 JOURNAL Submitted (04-MAR-1997) Hiroshi Ishiguro, Fujita Health University,  
 ICMS; 1-98, kutsukake-cho, Toyoake, Aichi 470-11, Japan  
 (E-mail:hishi@fujita-hu.ac.jp, Tel:0562-93-9393, Fax:0562-93-8831)  
 FEATURES Location/Qualifiers  
 source 1. .2188  
 /organism="Rattus sp."  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10118"  
 /cell\_line="PC12"  
 /cell\_type="pheochromocytoma"  
 CDS <1. .1926  
 /note="neural- and thymus-derived activator for ErbB  
 kinases (NTAK); a member of the epidermal growth factor  
 (EGF) family"  
 /codon\_start=1  
 /product="NTAK alpha2-1p"  
 /protein\_id="BAA23348.1"  
 /db\_xref="GI:2605479"  
 /translation="FFFFKTAFAVPDPNGKNIKKEVGKILCTDCATRPKLKKMKSQTG  
 EVGEKQSLKCEAAAGNPQPSYRWFKDGKELNRSRDIRIKYGNRKNRSLQFNKVKVED  
 AGEYVCEAENILGKDTVGRHLVNSVSTTLSSWSGHARKCNETAKSYCVNGGVCCYIE  
 GINQLSCKCPNGFFGQRCLEKLPLRLYMPDPKQKAEELYQKRVLTITGICVALLVGI  
 VCVVAYCKTKKQRRQMHHHLRQNMCPAHQNRSLANGPSHPRLDPEEIQMADYISKNP  
 ATDHVIRREAETTFSGSHSCSPSHHCSTATPTSSHRHESHTWSLERSESLSQSGI  
 MLSSVGTSKCNSPACVEARARRAAAYSQEERRRAAMPYPYHDSIDSLRDSPHSERYVSA  
 LTTPARLSPVDFHYSLATQVPTFEITSPNSAHAVSLPPAAPISYRLAEQQPLLRLHPAP  
 PGPGPGPGADMQRSYDSYYPAGPGPRRGACALGGSGLPASPFRIPEDEYETQ  
 ECAPPPPPRPRTRGASRRTSAGPRRWRRSRLNGLAAQRARAARDSLSLSSGSGCGSAS  
 ASDDDADDADGALAAESTPFLGLRAAHDALRSDSPPLCPAADSRTYYSLDSHSTRASS  
 RHSRGPPTRAKQDSGPL"

# ORIGIN

Query Match 47.4%; Score 425; DB 10; Length 2188;  
 Best Local Similarity 87.0%; Pred. No. 4.2e-71;  
 Matches 467; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 337 TTAGTCTTTAAGACGGCCTTTGCCCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAG 396  
 || | ||||| ||||| |||| | ||||| ||||| |||||  
 Db 4 TTTT TTTT TTAAGACAGCCTTTGCCCCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAG 63

QY 397 GTGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGC 456  
 |||||  
 Db 64 GTGGGCAAGATCCTGTGCACTGACTGCGCAACCCGGCCCAAGCTGAAGAAGATGAAGAGT 123  
 QY 457 CAGACGGGACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCC 516  
 ||||| ||| ||||| ||||| ||||| ||||| ||||| || || || || ||  
 Db 124 CAGACAGGAGAGGTGGGCGAGAAGCAGTCGCTCAAGTGTGAGGCGGCGGGGAACCCC 183  
 QY 517 CAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGC 576  
 ||||| ||||| || ||||| ||||| ||||| ||||| || || |||||  
 Db 184 CAGCCCTCCTATCGATGGTTCAAGGACGGCAAGGAGCTCAACCGGAGTCGTGACATTTCGC 243  
 QY 577 ATCAAATATGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAG 636  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 244 ATCAAGTATGGCAACGGCAGAAAGAACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAG 303  
 QY 637 GACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGC 696  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| || |||||  
 Db 304 GACGCTGGAGAGTACGTCTGTGAGGCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGC 363  
 QY 697 CGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCTCCTGGTCCGGGGCACGCCCGGAAG 756  
 ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 364 CGGCTCCATGTCAACAGTGTGAGCACCCTGTCTCCTGGTCCGGGGCACGCCCGGAAG 423  
 QY 757 TGCAACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGC 816  
 ||||| ||||| ||||| ||||| || || ||||| ||||| ||||| ||||| |||||  
 Db 424 TGCAATGAGACAGCCAAGTCCTACTGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGC 483  
 QY 817 ATCAACCAGCTCTCCTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 484 ATCAACCAACTCTCCTGCAAATGTCAAACGGATTCTTCGGACAGAGATGTTTGGAG 540

# RESULT 15

AC094080

LOCUS AC094080 118504 bp DNA linear PRI 27-MAR-2002

DEFINITION Homo sapiens chromosome 5 clone CTB-77K22, complete sequence.

ACCESSION AC094080

VERSION AC094080.4 GI:19747152

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 118504)

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 118504)

AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission

JOURNAL Submitted (14-SEP-2001) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 3 (bases 1 to 118504)

AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission

JOURNAL Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint

REFERENCE Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 4 (bases 1 to 118504)  
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-MAR-2002) DOE Joint Genome Institute, 2800 Mitchell  
 Drive, Walnut Creek, CA 94598, USA  
 COMMENT On Mar 27, 2002 this sequence version replaced gi:19224838.  
 Draft Sequence Produced by DOE Joint Genome Institute  
 www.jgi.doe.gov  
 Finishing Completed at Stanford Human Genome Center  
 www-shgc.stanford.edu  
 Quality: Phrap Quality >=40 99.7% of Sequence;  
 Estimated Total Number of Errors is 0.5.

FEATURES Location/Qualifiers  
 source 1. .118504  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 /clone="CTB-77K22"

# ORIGIN

Query Match 47.4%; Score 424.8; DB 9; Length 118504;  
 Best Local Similarity 94.2%; Pred. No. 3.5e-71;  
 Matches 452; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC	60
Db	80890	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC	80949
Qy	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	80950	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	81009
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	81010	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	81069
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	81070	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	81129
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	81130	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	81189
Qy	301	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	81190	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	81249
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	81250	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	81309
Qy	421	TGC-GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAA	479
Db	81310	TGCGGTGAGTCGCCCCCTCCCTTTGCTGGAGAAAGGGGGAGGGGCGAGGTGGTGGAGAA	81369

Search completed: August 15, 2004, 08:02:13  
Job time : 3832.5 secs

OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 02:37:58 ; Search time 430.237 Seconds  
(without alignments)  
8857.048 Million cell updates/sec

Title: US-09-864-675-3  
Perfect score: 897  
Sequence: 1 atgaggcgcgacccggcccc.....caatggtcaacttctcctaa 897

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002s:\*
- 7: geneseqn2003as:\*
- 8: geneseqn2003bs:\*
- 9: geneseqn2003cs:\*
- 10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	%					Description
	No.	Score	Query Match Length	DB	ID	
1	897	100.0	897	6	AAS18020	Aas18020 Human cDN
2	849	94.6	994	6	AAS18019	Aas18019 Human cDN
3	835.4	93.1	1884	2	AAV17814	Aav17814 Homo sapi
4	784	87.4	3441	2	AAT87922	Aat87922 Rat cereb
5	738.6	82.3	1803	2	AAT87923	Aat87923 Rat cereb
6	737	82.2	3076	2	AAV43674	Aav43674 Receptor
7	723.4	80.6	1863	7	ABS56035	Abs56035 cDNA enco

8	427.8	47.7	2268	2	AAV17816	Aav17816	Homo sapi
9	426.8	47.6	1474	7	ABS56036	Abs56036	cDNA enco
10	426.8	47.6	2266	7	ABS56045	Abs56045	cDNA enco
11	426.2	47.5	1476	2	AAV17815	Aav17815	Homo sapi
12	424	47.3	1054	6	ABL40993	Abl40993	Human neu
13	405.4	45.2	1607	2	AAV17813	Aav17813	Mus muscu
14	402.2	44.8	2467	2	AAV17812	Aav17812	Mus muscu
15	393.4	43.9	1561	7	ABS56034	Abs56034	cDNA enco
16	393.4	43.9	2442	7	ABS56033	Abs56033	cDNA enco
17	256	28.5	667	2	AAT87924	Aat87924	Human cer
18	173	19.3	419	6	ABL40994	Abl40994	Human neu
19	124.6	13.9	480	6	ABL40995	Abl40995	Human neu
20	124	13.8	350	6	ABL40996	Abl40996	Human neu
21	85.6	9.5	1986	2	AAZ32061	Aaz32061	Human MET
22	84	9.4	1986	5	AAC90318	Aac90318	L12260 cD
23	84	9.4	2003	2	AAT30995	Aat30995	Glial gro
24	84	9.4	2003	2	AAT48090	Aat48090	Human neu
25	84	9.4	2003	2	AAT06739	Aat06739	Glial gro
26	84	9.4	2003	2	AAZ32062	Aaz32062	Human MET
27	84	9.4	2003	5	AAC90319	Aac90319	I36352 cD
28	83.4	9.3	1108	2	AAQ30671	Aaq30671	GGF2BPP3.
29	83.4	9.3	1108	2	AAQ62850	Aaq62850	GGF-II cD
30	83.4	9.3	1108	2	AAQ58304	Aaq58304	GGF-II cD
31	83.4	9.3	1108	2	AAQ74886	Aaq74886	Putative
32	83.4	9.3	1108	2	AAT30998	Aat30998	Bovine gl
33	83.4	9.3	1108	2	AAT48080	Aat48080	Bovine ne
34	83.4	9.3	1108	2	AAT06704	Aat06704	Bovine gl
35	82.4	9.2	2003	2	AAQ62838	Aaq62838	GGF2HBS5.
36	82.4	9.2	2003	2	AAQ58329	Aaq58329	GGF2HBS5.
37	78.2	8.7	1027	5	AAF80062	Aaf80062	Nucleotid
38	78.2	8.7	3086	5	AAF80059	Aaf80059	Nucleotid
39	76.8	8.6	1140	2	AAQ62840	Aaq62840	GGF2BPP2.
40	76.8	8.6	1140	2	AAQ58321	Aaq58321	GGF2BPP2.
41	76.8	8.6	1140	2	AAQ74912	Aaq74912	Bovine gl
42	76.8	8.6	1140	2	AAT31001	Aat31001	Glial gro
43	76.8	8.6	1140	2	AAT48088	Aat48088	Human neu
44	76.8	8.6	1140	2	AAT06731	Aat06731	BPP2 glia
45	76.8	8.6	1140	2	AAX81201	Aax81201	Nucleotid

# ALIGNMENTS

## RESULT 1

AAS18020

ID AAS18020 standard; cDNA; 897 BP.

XX

AC AAS18020;

XX

DT 12-MAR-2002 (first entry)

XX

DE Human cDNA encoding Neuregulin-2beta, NRG-2beta.

XX

KW Human; ss; neuregulin-2; NRG-2alpha; NRG-2beta; mitogenesis;

KW cell survival; cell growth; cell differentiation; erbB receptor;

KW cardiomyopathy; ischaemic damage; cardiac trauma; heart failure;

KW atherosclerosis; vascular lesion; vascular hypertension;

KW degenerative congenital vascular disease; myasthenia gravis;  
KW neurodegenerative disorder; peripheral neuropathy;  
KW sensory nerve fiber neuropathy; motor fiber neuropathy;  
KW sensory nerve fiber neuropathy; multiple sclerosis;  
KW amyotrophic lateral sclerosis; spinal muscular atrophy; nerve injury;  
KW Alzheimer's disease; Parkinson's disease; cerebellar ataxia;  
KW spinal cord injury; tumour; neurofibromatosis; transgenic animal.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers  
FT CDS 1. .897  
FT /\*tag= a  
FT /product= "NRG-2beta"

XX

PN WO200189568-A1.

XX

PD 29-NOV-2001.

XX

PF 23-MAY-2001; 2001WO-US016896.

XX

PR 23-MAY-2000; 2000US-0206495P.

XX

PA (CENE-) CENES PHARM INC.

XX

PI Marchionni MA;

XX

DR WPI; 2002-097612/13.

DR

P-PSDB; AAU11636.

XX

PT Neuregulin-2 polypeptide and polynucleotide useful for treating multiple  
PT sclerosis, spinal muscular atrophy, nerve injury, Alzheimer's disease, by  
PT increasing mitogenesis, survival, growth or differentiation of a cell.

XX

PS Claim 57; Fig 8; 79pp; English.

XX

CC The invention relates to a substantially pure neuregulin (NRG)-2  
CC polypeptide comprising or consisting of a sequence for human NRG-2alpha  
CC or NRG-2beta (clone 2b7) and the polynucleotides encoding the. Also  
CC included are a vector expressing the protein, a host cell comprising the  
CC vector, a transgenic non-human animal transformed with the vector or  
CC having a knockout mutation in one or both NRG-2 alleles and an anti-NRG-2  
CC antibody. Analysis of mutations in NRG-2 in an individual is useful for  
CC diagnosing an increased likelihood of developing a NRG-2-related disease  
CC or condition in a test subject. NRG-2 is useful for increasing the  
CC mitogenesis, survival, growth or differentiation of a cell (e.g. a  
CC neuronal cell), where the cell expresses an erbB receptor. NRG-2 is  
CC useful for treating diseases and disorders such as cardiomyopathy  
CC (preferably degenerative congenital disease), ischaemic damage, cardiac  
CC trauma or heart failure or which has a condition affecting smooth muscle  
CC which include atherosclerosis, vascular lesion, vascular hypertension,  
CC and degenerative congenital vascular disease, myasthenia gravis, a  
CC neurodegenerative disorder, peripheral neuropathy, a sensory nerve fiber  
CC neuropathy, a motor fiber and a sensory nerve fiber neuropathy, multiple  
CC sclerosis, amyotrophic lateral sclerosis, spinal muscular atrophy, nerve  
CC injury, Alzheimer's disease, Parkinson's disease, cerebellar ataxia, and  
CC spinal cord injury. The antibody is useful for treatment of a tumour



CC comprising inhibiting proliferation of a tumour cell preferably a glial  
CC tumour cell, for treating of neurofibromatosis by inhibiting glial cell  
CC mitogenesis. The present sequence encodes NRG-2beta

XX

SQ Sequence 897 BP; 200 A; 261 C; 282 G; 154 T; 0 U; 0 Other;

Query Match 100.0%; Score 897; DB 6; Length 897;

Best Local Similarity 100.0%; Pred. No. 1.5e-195;

Matches 897; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60
        |||
Db      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60

Qy     61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
        |||
Db     61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120

Qy    121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
        |||
Db    121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180

Qy    181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
        |||
Db    181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240

Qy    241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
        |||
Db    241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300

Qy    301 CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
        |||
Db    301 CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360

Qy    361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
        |||
Db    361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420

Qy    421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
        |||
Db    421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480

Qy    481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
        |||
Db    481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540

Qy    541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 600
        |||
Db    541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 600

Qy    601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
        |||
Db    601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660

Qy    661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
        |||
Db    661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
```

Qy 721 ACCACCCTGTCATCCTGGTCGGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780  
 |||  
 Db 721 ACCACCCTGTCATCCTGGTCGGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780  
 Qy 781 TCGCTCAATGGAGGCGTCTGCTACTACATCGAGGGGCATCAACCAGCTCTCCTGCAAGTGT 840  
 |||  
 Db 781 TCGCTCAATGGAGGCGTCTGCTACTACATCGAGGGGCATCAACCAGCTCTCCTGCAAGTGT 840  
 Qy 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGGTCAACTTCTCCTAA 897  
 |||  
 Db 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGGTCAACTTCTCCTAA 897

RESULT 2

AAS18019

ID AAS18019 standard; cDNA; 994 BP.

XX

AC AAS18019;

XX

DT 12-MAR-2002 (first entry)

XX

DE Human cDNA encoding Neuregulin-2alpha, NRG-2alpha.

XX

KW Human; ss; neuregulin-2; NRG-2alpha; NRG-2beta; mitogenesis;  
 KW cell survival; cell growth; cell differentiation; erbB receptor;  
 KW cardiomyopathy; ischaemic damage; cardiac trauma; heart failure;  
 KW atherosclerosis; vascular lesion; vascular hypertension;  
 KW degenerative congenital vascular disease; myasthenia gravis;  
 KW neurodegenerative disorder; peripheral neuropathy;  
 KW sensory nerve fiber neuropathy; motor fiber neuropathy;  
 KW sensory nerve fiber neuropathy; multiple sclerosis;  
 KW amyotrophic lateral sclerosis; spinal muscular atrophy; nerve injury;  
 KW Alzheimer's disease; Parkinson's disease; cerebellar ataxia;  
 KW spinal cord injury; tumour; neurofibromatosis; transgenic animal.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
FT	CDS	1. .993
FT		/*tag= a
FT		/product= "NRG-2alpha"

XX

PN WO200189568-A1.

XX

PD 29-NOV-2001.

XX

PF 23-MAY-2001; 2001WO-US016896.

XX

PR 23-MAY-2000; 2000US-0206495P.

XX

PA (CENE-) CENES PHARM INC.

XX

PI Marchionni MA;

XX

DR WPI; 2002-097612/13.

DR P-PSDB; AAU11635.



Qy	301	CGCTACATCTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	301	CGCTACATCTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Qy	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAG	600
Db	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAG	600
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Qy	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT	840
Db	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT	840
Qy	841	CCTGTGGGATACACCGGGGACAGGTGTCAGCAG	873
Db	841	CCAAATGGATTCTTCGGACAGAGATGTTTGGAG	873

# RESULT 3

AAV17814

ID AAV17814 standard; cDNA; 1884 BP.

XX

AC AAV17814;

XX

DT 17-AUG-1998 (first entry)

XX

DE Homo sapiens don-1 gene splice variant.

XX

KW Murine; don-1 gene; melanoma; treatment; adenocarcinoma; epithelial cell;

KW proliferation; stimulation; treatment; tumours; skin; oesophagus; lung;

KW breast; liver; pancreas; colon; prostate; gastrointestinal tract; uterus;

KW wound healing; transmembrane; ss.

XX

OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 664. .1884  
 FT /\*tag= a  
 FT /note= "don-1 polypeptide"  
 XX  
 PN WO9807736-A1.  
 XX  
 PD 26-FEB-1998.  
 XX  
 PF 18-AUG-1997; 97WO-US014585.  
 XX  
 PR 19-AUG-1996; 96US-00699591.  
 PR 19-NOV-1996; 96US-00753007.  
 XX  
 PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.  
 XX  
 PI Gearing DP, Busfield SJ;  
 XX  
 DR WPI; 1998-169084/15.  
 DR P-PSDB; AAW48381.  
 XX  
 PT Mouse and human don-1 polypeptide(s) - useful for treatment of melanomas  
 PT and adenocarcinoma(s), and for wound healing.  
 XX  
 PS Claim 4; Fig 3; 121pp; English.  
 XX  
 CC The sequence is that of a human don-1 gene splice variant. Don-1  
 CC polypeptides stimulate proliferation of epithelial cells and thus are  
 CC implicated in melanomas and adenocarcinomas in which epithelial cells  
 CC proliferate out of control. Compounds that interfere with don-1 mediated  
 CC cell proliferation can be used in the treatment of tumours such as  
 CC melanomas and adenocarcinomas of the skin, oesophagus, lung, breast,  
 CC liver, pancreas, gastrointestinal tract, colon, prostate or uterus.  
 CC Alternatively, don-1 polypeptides can be used to stimulate epithelial  
 CC cell proliferation, e.g. for wound healing  
 XX  
 SQ Sequence 1884 BP; 426 A; 607 C; 560 G; 291 T; 0 U; 0 Other;

Query Match 93.1%; Score 835.4; DB 2; Length 1884;  
 Best Local Similarity 98.1%; Pred. No. 2.2e-181;  
 Matches 856; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

Qy 1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 218 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 277  
 Qy 61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 278 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 337  
 Qy 121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 338 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 397  
 Qy 181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240

Db	398	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	457
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	458	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	517
Qy	301	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	518	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	577
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	578	CCCCT-GATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC	636
Qy	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	637	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	696
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	697	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	756
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	757	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG	816
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	817	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	876
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	877	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	936
Qy	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	937	ACCACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	996
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT	840
Db	997	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT	1056
Qy	841	CCTGTGGGATACACCGGGGACAGGTGTCAGCAG	873
Db	1057	CCAAATGGATTCTTCGGACAGAGATGTTTGGAG	1089

RESULT 4

AAT87922

ID AAT87922 standard; cDNA; 3441 BP.

XX

AC AAT87922;

XX

DT 18-DEC-1997 (first entry)

XX

DE Rat cerebellum derived growth factor 1 cDNA.

XX  
 KW Rat; cerebellum derived growth factor; CDGF1; screening; binding;  
 KW modulation; erbB type receptor; identification; indication; risk;  
 KW proliferation; differentiation; induction; neuron; hyperplasia;  
 KW stem cell culture; intracerebral graft; alleviation; repair;  
 KW behavioural defect; nervous system; central; peripheral; nerve;  
 KW prosthesis; damage; entubulation; cell survival; treatment; injury;  
 KW trauma; ischaemia; ischemia; stroke; infection; disorder; inflammation;  
 KW neurodegeneration; disease; Parkinson's; Huntingdon's;  
 KW amyotrophic lateral sclerosis; sensory; retina;  
 KW spinocerebellar degeneration; multiple sclerosis; neoplasia;  
 KW amalignant glioma; medulloblastoma; neuroectodermal tumour; ds.  
 XX  
 OS Rattus rattus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 180. .2444  
 FT /\*tag= a  
 FT sig\_peptide 180. .248  
 FT /\*tag= b  
 FT mat\_peptide 249. .2441  
 FT /\*tag= c  
 FT /product= "cerebellum\_derived\_growth\_factor"  
 XX  
 PN WO9709425-A1.  
 XX  
 PD 13-MAR-1997.  
 XX  
 PF 09-SEP-1996; 96WO-US014484.  
 XX  
 PR 08-SEP-1995; 95US-00525864.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 PA (STRD ) UNIV LELAND S STANFORD.  
 XX  
 PI Chang H;  
 XX  
 DR WPI; 1997-192900/17.  
 DR P-PSDB; AAW27536.  
 XX  
 PT Rat and human cerebellum-derived growth factors - used in the treatment  
 PT of neuronal injury and proliferative disorders.  
 XX  
 PS Claim 17; Page 63-66; 94pp; English.  
 XX  
 CC The present sequence encodes rat cerebellum derived growth factor 1  
 CC (CDGF1), which can be used to screen for modulators of CDGF binding to  
 CC erbB type receptors. Identification of a modification or mutation in a  
 CC CDGF gene, or aberrant expression of a CDGF gene or levels of soluble  
 CC CDGF may be used to indicate the risk of unwanted cell proliferation or  
 CC differentiation. CDGF may be used to induce neuronal differentiation in  
 CC stem cell culture, and maintain the integrity of a terminally  
 CC differentiated neuronal cell culture, e.g. useful for intracerebral  
 CC grafting to alleviate behavioural defects. CDGF may also be used in nerve  
 CC prostheses to repair central and peripheral nerve damage, especially where  
 CC a crushed or severed axon is entubulated by a prosthetic. CDGF may also  
 CC be used to enhance neuronal cell survival in the central or peripheral

CC nervous system, to treat neurological conditions associated with nervous  
 CC system injury, e.g. traumatic, chemical or vasal injury and deficits such  
 CC as ischaemia resulting from stroke, infectious/inflammatory and tumour  
 CC induced injury, chronic neurodegenerative disease including Parkinson's  
 CC and Huntingdon's, amyotrophic lateral sclerosis, spinocerebellar  
 CC degeneration, chronic immunological disease of the nervous system  
 CC including multiple sclerosis, disorders of the sensory neurons and  
 CC degenerative diseases of the retina. CDGF may also be used to treat  
 CC neoplastic or hyperplastic transformations, particularly of the central  
 CC nervous system, e.g. amalignant gliomas, medulloblastomas and  
 CC neuroectodermal tumours

XX

SQ Sequence 3441 BP; 777 A; 1057 C; 1015 G; 592 T; 0 U; 0 Other;

Query Match 87.4%; Score 784; DB 2; Length 3441;

Best Local Similarity 92.2%; Pred. No. 1.4e-169;

Matches 826; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC	60
Db	180	ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC	239
Qy	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	240	TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	299
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	300	GGCAAGGTACAGGGACTGGCCCCGGCAGGCGTTCCAGCTCTAACAGCACCCGAGAGCCT	359
Qy	181	CCCGCCTCGGGTGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	360	CCCGCCTCGGGTGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	419
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	420	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG	479
Qy	301	CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	480	CGCTACATCTTTTTCCTGGAGCCACCGAGCAGCCCTTAGT'TTTAAGACAGCCTTTGCC	539
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	540	CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	599
Qy	421	TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	600	TGCGCAACCCGGCCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG	659
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	660	CAGTCGCTCAAGTGTGAGGCGGCGGGGAACCCCGCCCTCCTATCGATGGTTCAAG	719
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	720	GACGGCAAGGAGCTCAACCGGAGTCGTGACATTCGCATCAAGTATGGCAACGGCAGAAAG	779





PD 13-MAR-1997.  
 XX  
 PF 09-SEP-1996; 96WO-US014484.  
 XX  
 PR 08-SEP-1995; 95US-00525864.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 PA (STRD ) UNIV LELAND S STANFORD.  
 XX  
 PI Chang H;  
 XX  
 DR WPI; 1997-192900/17.  
 DR P-PSDB; AAW27537.  
 XX  
 PT Rat and human cerebellum-derived growth factors - used in the treatment  
 PT of neuronal injury and proliferative disorders.  
 XX  
 PS Claim 17; Page 70-74; 94pp; English.  
 XX  
 CC The present sequence encodes rat cerebellum derived growth factor 2  
 CC (CDGF2), which can be used to screen for modulators of CDGF binding to  
 CC erbB type receptors. Identification of a modification or mutation in a  
 CC CDGF gene, or aberrant expression of a CDGF gene or levels of soluble  
 CC CDGF may be used to indicate the risk of unwanted cell proliferation or  
 CC differentiation. CDGF may be used to induce neuronal differentiation in  
 CC stem cell culture, and maintain the integrity of a terminally  
 CC differentiated neuronal cell culture, e.g. useful for intracerebral  
 CC grafting to alleviate behavioural defects. CDGF may also be used in nerve  
 CC prostheses to repair central and peripheral nerve damage, especially where  
 CC a crushed or severed axon is entubulated by a prosthetic. CDGF may also  
 CC be used to enhance neuronal cell survival in the central or peripheral  
 CC nervous system, to treat neurological conditions associated with nervous  
 CC system injury, e.g. traumatic, chemical or vasal injury and deficits such  
 CC as ischaemia resulting from stroke, infectious/inflammatory and tumour  
 CC induced injury, chronic neurodegenerative disease including Parkinson's  
 CC and Huntingdon's, amyotrophic lateral sclerosis, spinocerebellar  
 CC degeneration, chronic immunological disease of the nervous system  
 CC including multiple sclerosis, disorders of the sensory neurons and  
 CC degenerative diseases of the retina. CDGF may also be used to treat  
 CC neoplastic or hyperplastic transformations, particularly of the central  
 CC nervous system, e.g. amalignant gliomas, medulloblastomas and  
 CC neuroectodermal tumours  
 XX  
 SQ Sequence 1803 BP; 408 A; 549 C; 537 G; 309 T; 0 U; 0 Other;

Query Match 82.3%; Score 738.6; DB 2; Length 1803;  
 Best Local Similarity 90.4%; Pred. No. 3e-159;  
 Matches 789; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC	60
Db	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC	60
Qy	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	61	TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120

Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	121	GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT	180
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	181	CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG	300
Qy	301	CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	301	CGCTACATCTTTTTCTGGAGCCACCGAGCAGCCCTTAGTTTTTAAGACAGCCTTTGCC	360
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	361	CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Qy	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	421	TGCGCAACCCGGCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG	480
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	481	CAGTCGCTCAAGTGTGAGGCGGCGGGGAACCCAGCCCTCCTATCGATGGTTCAAG	540
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	541	GACGGCAAGGAGCTCAACCGGAGTCGTGACATTCGCATCAAGTATGGCAACGGCAGAAAG	600
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	601	AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG	660
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	661	GCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGTGAGC	720
Qy	721	ACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	721	ACCACTCTGTCGTCTGGTCGGGGCACGCCCGGAAGTGCAATGAGACAGCCAAGTCCTAC	780
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT	840
Db	781	TGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAATGT	840
Qy	841	CCTGTGGGATACACCGGGGACAGGTGTCAGCAG	873
Db	841	CCAAACGGATTCTTCGGACAGAGATGTTTGGAG	873

RESULT 6

AAV43674

ID AAV43674 standard; cDNA; 3076 BP.

XX

AC AAV43674;  
 XX  
 DT 29-SEP-1998 (first entry)  
 XX  
 DE Receptor type tyrosine kinase ErbB ligand encoding cDNA.  
 XX  
 KW Receptor type tyrosine kinase ErbB; ligand; diagnostic agent;  
 KW nervous disease; cancer; ss.  
 XX  
 OS Rattus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 232..2814  
 FT /\*tag= a  
 FT /product= "ligand of receptor type tyrosine kinase ErbB"  
 XX  
 PN JP10179166-A.  
 XX  
 PD 07-JUL-1998.  
 XX  
 PF 25-DEC-1996; 96JP-00356998.  
 XX  
 PR 25-DEC-1996; 96JP-00356998.  
 XX  
 PA (HIGA/) HIGASHIYAMA S.  
 XX  
 DR WPI; 1998-430952/37.  
 DR P-PSDB; AAW63700.  
 XX  
 PT Gene coding the ligand of the tyrosine kinase ErbB receptor - useful for  
 PT diagnosing and treating nervous diseases and cancer.  
 XX  
 PS Example; Page 9-13; 17pp; Japanese.  
 XX  
 CC This cDNA encodes the ligand of receptor type tyrosine kinase ErbB. A  
 CC prokaryotic or eukaryotic host cell transformed by a recombinant vector  
 CC containing the encoding DNA can be used for the recombinant production of  
 CC the protein. The invention provides a method for inhibiting the formation  
 CC of the ligand of receptor type tyrosine kinase ErbB in an animal using an  
 CC antibody recognizing the protein. The ligand of the tyrosine kinase ErbB  
 CC receptor and associated materials can be used for treating or diagnosing  
 CC nervous diseases and cancers  
 XX  
 SQ Sequence 3076 BP; 673 A; 996 C; 944 G; 463 T; 0 U; 0 Other;

Query Match 82.2%; Score 737; DB 2; Length 3076;  
 Best Local Similarity 90.3%; Pred. No. 7.9e-159;  
 Matches 788; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy 1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 60  
 ||||||||||||||||||||||||| ||| ||||||||||||||||||| ||||||||  
 Db 556 ATGAGGCGCGACCCGGCCCCCGGCTCCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 615  
 Qy 61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120  
 ||||||||||||||||||||| |||||||||||||||||||||||||||||  
 Db 616 TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 675

Qy 121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180  
 |||||  
 Db 676 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT 735

Qy 181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240  
 |||||  
 Db 736 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 795

Qy 241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300  
 |||||  
 Db 796 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG 855

Qy 301 CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360  
 |||||  
 Db 856 CGCTACATCTTTTTCTGGAGCCACCGAGCAGCCCTTAGTTTTTAAGACAGCCTTTGCC 915

Qy 361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420  
 || |||||  
 Db 916 CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 975

Qy 421 TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480  
 |||||  
 Db 976 TGCGCAACCCGGCCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG 1035

Qy 481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540  
 || |||||  
 Db 1036 CAGTCGCTCAAGTGTGAGGCGGCGGGGAACCCCCAGCCCTCCTATCGATGGTTCAAG 1095

Qy 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG 600  
 || |||||  
 Db 1096 GACGGCAAGGAGCTCAACCGGAGTCGTGACATTCGCATCAAGTATGGCAACGGCAGAAAG 1155

Qy 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660  
 |||||  
 Db 1156 AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG 1215

Qy 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720  
 || |||||  
 Db 1216 GCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGTGAGC 1275

Qy 721 ACCACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780  
 |||||  
 Db 1276 ACCACTCTGTCTCCTGGTCGGGGCACGCCCGGAAGTGCAATGAGACAGCCAAGTCCTAC 1335

Qy 781 TCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840  
 || |||||  
 Db 1336 TGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAAATGT 1395

Qy 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873  
 || |||||  
 Db 1396 CCAAACGGATTCTTCGGACAGAGATGTTTGGAG 1428

RESULT 7

ABS56035

ID ABS56035 standard; cDNA; 1863 BP.

XX

AC ABS56035;  
 XX  
 DT 14-JAN-2003 (first entry)  
 XX  
 DE cDNA encoding human membrane-bound splice variant of Don-1.  
 XX  
 KW Human; Don-1; epidermal growth factor; EGF; neuregulin;  
 KW glycoprotein ligand; cell proliferation; cell proliferative disorder;  
 KW carcinoma; adenocarcinoma cell; myeloma; cell differentiation;  
 KW cell survival; epithelial cell; wound healing; tumour formation; brain;  
 KW vulnerary; cytostatic; gene therapy; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 643..1863  
 FT /\*tag= a  
 FT /partial  
 FT /product= "Membrane-bound splice variant of Don-1"  
 FT /note= "This sequence lacks a stop codon"  
 XX  
 PN US2002127594-A1.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 12-MAR-2002; 2002US-00096241.  
 XX  
 PR 22-JUN-2000; 2000US-00599789.  
 XX  
 PA (GEAR/) GEARING D P.  
 PA (BUSF/) BUSFIELD S J.  
 XX  
 PI Gearing DP, Busfield SJ;  
 XX  
 DR WPI; 2003-039584/03.  
 DR P-PSDB; ABG71638.  
 XX  
 PT Novel Don-1 polypeptide useful for stimulating proliferation of cells,  
 PT for identifying proteins that interact with Don-1, and for regulating  
 PT tumor formation and progression in brain.  
 XX  
 PS Claim 4; Fig 3; 66pp; English.  
 XX  
 CC The present invention relates to the isolation of a novel gene called Don  
 CC -1, and alternate splice variants of Don-1, which are related to  
 CC epidermal growth factors (EGF) such as neuregulins. Don-1 polypeptides  
 CC are glycoprotein ligands. Both murine and human Don-1 sequences are  
 CC cloned. The mouse Don-1 gene maps to chromosome 18. Don-1 polypeptides  
 CC are useful for stimulating proliferation of a cell. Antibodies to Don-1  
 CC polypeptides are useful for detecting Don-1 in a sample. The Don-1  
 CC polypeptides are useful for treating and diagnosing cell proliferative  
 CC disorders and play a role in the proliferation of carcinomas e.g.  
 CC adenocarcinoma, myeloma, in cell differentiation, proliferation and  
 CC survival. The polypeptides are also useful for inhibiting proliferation  
 CC of adenocarcinoma cells, for stimulating the proliferation of cells such  
 CC as epithelial cells to promote wound healing, for identifying proteins  
 CC that interact with Don-1, and for regulating tumour formation and

CC progression in the brain. The polynucleotide sequences encoding Don-1 may  
CC be used in gene therapy. The present sequence encodes human membrane-  
CC bound splice variant of Don-1

XX

SQ Sequence 1863 BP; 422 A; 602 C; 553 G; 286 T; 0 U; 0 Other;

Query Match 80.6%; Score 723.4; DB 7; Length 1863;  
Best Local Similarity 96.2%; Pred. No. 9.2e-156;  
Matches 840; Conservative 0; Mismatches 16; Indels 17; Gaps 9;

```
Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60
      |||
Db     213 ATGAGGCGCGACCCGGCCCCC--CTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 270

Qy      61 TACTCGCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
      |||
Db     271 TACTCGCCAGCCTCAAGTCA--GCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 328

Qy     121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
      |||
Db     329 GGCAAGGTACAGGGGCTGGT--CAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 386

Qy     181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
      |||
Db     387 CCCGCCTCGGGTCGGGTGGCG--GGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 444

Qy     241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
      |||
Db     445 GGGCTGCAGCGCGAGCAGGTG--CAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 502

Qy     301 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
      |||
Db     503 CGCTACATCTTTTTCCTGGAG--CACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 560

Qy     361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
      |||
Db     561 CCCCT-GATACCAACGGCAAAA--CTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC 617

Qy     421 TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
      |||
Db     618 TGCGCCACCCGGCCCCAAGTTGA--AAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 675

Qy     481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540
      |||
Db     676 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 735

Qy     541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 600
      |||
Db     736 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG 795

Qy     601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
      |||
Db     796 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 855

Qy     661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
      |||
Db     856 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 915
```

QY 721 ACCACCCTGTCATCCTGGTCGGGGGCACGCCCAGGAGTCAACGAGACAGCCAAGTCCTAT 780  
 |||  
 Db 916 ACCACCCTGTCATCCTGGTCGGGGGCACGCCCAGGAGTCAACGAGACAGCCAAGTCCTAT 975

QY 781 TGCCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840  
 |||  
 Db 976 TGCCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAAATGT 1035

QY 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873  
 || ||| | || | || || ||  
 Db 1036 CCAATGGATTCTTCGGACAGAGATGTTTGGAG 1068

# RESULT 8

AAV17816

ID AAV17816 standard; cDNA; 2268 BP.

XX

AC AAV17816;

XX

DT 17-AUG-1998 (first entry)

XX

DE Homo sapiens don-1 gene splice variant.

XX

KW Murine; don-1 gene; melanoma; treatment; adenocarcinoma; epithelial cell;  
 KW proliferation; stimulation; treatment; tumours; skin; oesophagus; lung;  
 KW breast; liver; pancreas; colon; prostate; gastrointestinal tract; uterus;  
 KW wound healing; transmembrane; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 69. .2012

FT /\*tag= a

FT /note= "don-1 polypeptide"

XX

PN WO9807736-A1.

XX

PD 26-FEB-1998.

XX

PF 18-AUG-1997; 97WO-US014585.

XX

PR 19-AUG-1996; 96US-00699591.

PR 19-NOV-1996; 96US-00753007.

XX

PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.

XX

PI Gearing DP, Busfield SJ;

XX

DR WPI; 1998-169084/15.

DR P-PSDB; AAW48383.

XX

PT Mouse and human don-1 polypeptide(s) - useful for treatment of melanomas  
 PT and adenocarcinoma(s), and for wound healing.

XX

PS Claim 4; Fig 7; 121pp; English.

XX



CC The sequence is that of a human don-1 gene splice variant. Don-1  
CC polypeptides stimulate proliferation of epithelial cells and thus are  
CC implicated in melanomas and adenocarcinomas in which epithelial cells  
CC proliferate out of control. Compounds that interfere with don-1 mediated  
CC cell proliferation can be used in the treatment of tumours such as  
CC melanomas and adenocarcinomas of the skin, oesophagus, lung, breast,  
CC liver, pancreas, gastrointestinal tract, colon, prostate or uterus.  
CC Alternatively, don-1 polypeptides can be used to stimulate epithelial  
CC cell proliferation, e.g. for wound healing

XX

SQ Sequence 2268 BP; 502 A; 735 C; 700 G; 331 T; 0 U; 0 Other;

Query Match 47.7%; Score 427.8; DB 2; Length 2268;  
Best Local Similarity 89.8%; Pred. No. 4.1e-88;  
Matches 459; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```
Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
      || || | || | | | | || | | | | | | | | | |
Db      98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
      |||||
Db      158 AGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
      |||||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 602
      |||||
Db      278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAGAA 337

Qy      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
      |||||
Db      338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

Qy      663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
      |||||
Db      398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

Qy      723 CACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
      |||||
Db      458 CACCCTGTCATCCTGGTCGGGGCACGCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517

Qy      783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842
      |||||
Db      518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAATGTCC 577

Qy      843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873
      |||| | || | || | ||
Db      578 AAATGGATTCTTCGGACAGAGATGTTTGGAG 608
```

RESULT 9

ABS56036

ID ABS56036 standard; cDNA; 1474 BP.

XX

AC ABS56036;  
 XX  
 DT 14-JAN-2003 (first entry)  
 XX  
 DE cDNA encoding human second splice variant of Don-1.  
 XX  
 KW Human; Don-1; epidermal growth factor; EGF; neuregulin;  
 KW glycoprotein ligand; cell proliferation; cell proliferative disorder;  
 KW carcinoma; adenocarcinoma cell; myeloma; cell differentiation;  
 KW cell survival; epithelial cell; wound healing; tumour formation; brain;  
 KW vulnerary; cytostatic; gene therapy; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 68. .1473  
 FT /\*tag= a  
 FT /partial  
 FT /product= "Second splice variant of Don-1"  
 FT /note= "This sequence lacks a stop codon"  
 FT /transl\_except= (pos:107. .108, aa:Lys)  
 FT /note= "This codon has an apparent 1 nucleotide deletion  
 FT which alters the reading frame"  
 XX  
 PN US2002127594-A1.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 12-MAR-2002; 2002US-00096241.  
 XX  
 PR 22-JUN-2000; 2000US-00599789.  
 XX  
 PA (GEAR/) GEARING D P.  
 PA (BUSF/) BUSFIELD S J.  
 XX  
 PI Gearing DP, Busfield SJ;  
 XX  
 DR WPI; 2003-039584/03.  
 DR P-PSDB; ABG71639.  
 XX  
 PT Novel Don-1 polypeptide useful for stimulating proliferation of cells,  
 PT for identifying proteins that interact with Don-1, and for regulating  
 PT tumor formation and progression in brain.  
 XX  
 PS Claim 4; Fig 4; 66pp; English.  
 XX  
 CC The present invention relates to the isolation of a novel gene called Don  
 CC -1, and alternate splice variants of Don-1, which are related to  
 CC epidermal growth factors (EGF) such as neuregulins. Don-1 polypeptides  
 CC are glycoprotein ligands. Both murine and human Don-1 sequences are  
 CC cloned. The mouse Don-1 gene maps to chromosome 18. Don-1 polypeptides  
 CC are useful for stimulating proliferation of a cell. Antibodies to Don-1  
 CC polypeptides are useful for detecting Don-1 in a sample. The Don-1  
 CC polypeptides are useful for treating and diagnosing cell proliferative  
 CC disorders and play a role in the proliferation of carcinomas e.g.  
 CC adenocarcinoma, myeloma, in cell differentiation, proliferation and  
 CC survival. The polypeptides are also useful for inhibiting proliferation



XX  
 DE cDNA encoding human third splice variant of Don-1.  
 XX  
 KW Human; Don-1; epidermal growth factor; EGF; neuregulin;  
 KW glycoprotein ligand; cell proliferation; cell proliferative disorder;  
 KW carcinoma; adenocarcinoma cell; myeloma; cell differentiation;  
 KW cell survival; epithelial cell; wound healing; tumour formation; brain;  
 KW vulnerary; cytostatic; gene therapy; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 68..2010  
 FT /\*tag= a  
 FT /product= "Third splice variant of Don-1"  
 FT /transl\_except= (pos:107..108, aa:Lys)  
 FT /note= "This codon has an apparent 1 nucleotide deletion  
 FT which alters the reading frame"  
 FT /transl\_except= (pos:994..996, aa:Thr)  
 XX  
 PN US2002127594-A1.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 12-MAR-2002; 2002US-00096241.  
 XX  
 PR 22-JUN-2000; 2000US-00599789.  
 XX  
 PA (GEAR/) GEARING D P.  
 PA (BUSF/) BUSFIELD S J.  
 XX  
 PI Gearing DP, Busfield SJ;  
 XX  
 DR WPI; 2003-039584/03.  
 DR P-PSDB; ABG71644.  
 XX  
 PT Novel Don-1 polypeptide useful for stimulating proliferation of cells,  
 PT for identifying proteins that interact with Don-1, and for regulating  
 PT tumor formation and progression in brain.  
 XX  
 PS Claim 4; Fig 7; 66pp; English.  
 XX  
 CC The present invention relates to the isolation of a novel gene called Don  
 CC -1, and alternate splice variants of Don-1, which are related to  
 CC epidermal growth factors (EGF) such as neuregulins. Don-1 polypeptides  
 CC are glycoprotein ligands. Both murine and human Don-1 sequences are  
 CC cloned. The mouse Don-1 gene maps to chromosome 18. Don-1 polypeptides  
 CC are useful for stimulating proliferation of a cell. Antibodies to Don-1  
 CC polypeptides are useful for detecting Don-1 in a sample. The Don-1  
 CC polypeptides are useful for treating and diagnosing cell proliferative  
 CC disorders and play a role in the proliferation of carcinomas e.g.  
 CC adenocarcinoma, myeloma, in cell differentiation, proliferation and  
 CC survival. The polypeptides are also useful for inhibiting proliferation  
 CC of adenocarcinoma cells, for stimulating the proliferation of cells such  
 CC as epithelial cells to promote wound healing, for identifying proteins  
 CC that interact with Don-1, and for regulating tumour formation and  
 CC progression in the brain. The polynucleotide sequences encoding Don-1 may

CC be used in gene therapy. The present sequence encodes human third splice  
CC variant of Don-1

Query Match 47.6%; Score 426.8; DB 7; Length 2266;  
Best Local Similarity 92.4%; Pred. No. 6.9e-88;  
Matches 449; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Best Local Similarity 92.4%; Pred. No. 6.9e-88;

Matches 449; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy	388	AAGAAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCCGGCCCAAGTTGAAGAAG	447
Db	121	AGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAAAGCCACCCGGCCCAAGTTGAAGAAG	180
Qy	448	ATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCC	507
Db	181	ATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGAAGTGTGAGGCAGCAGCC	240
Qy	508	GGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGA	567
Db	241	GGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGGAGCTCAACCGCAGCCGA	300
Qy	568	GACATTGCGATCAAATATGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTG	627
Db	301	GACATTGCGATCAAATATGGCAACGGCAGAAAGAACTCACGACTACAGTTCAACAAGGTG	360
Qy	628	AAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACC	687
Db	361	AAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACATCCTGGGGAAGGACACC	420
Qy	688	GTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCATCCTGGTCGGGGCAC	747
Db	421	GTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGTCATCCTGGTCGGGGCAC	480
Qy	748	GCCCCGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTAC	807
Db	481	GCCCCGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATGGAGGCGTCTGCTACTAC	540
Qy	808	ATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGT	867
Db	541	ATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCTGTGGGATACACCGGGGACAGGTGT	600
Qy	868	CAGCAG	873
Db	601	TTGGAG	606

RESULT 11

ID    AAV17815 standard; cDNA; 1476 BP.

XX

XX

DT

DE

XX

KW proliferation; stimulation; treatment; tumours; skin; oesophagus; lung;  
KW breast; liver; pancreas; colon; prostate; gastrointestinal tract; uterus;  
KW wound healing; transmembrane; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 69. .1475  
FT /\*tag= a  
FT /note= "don-1 polypeptide"  
XX  
PN WO9807736-A1.  
XX  
PD 26-FEB-1998.  
XX  
PF 18-AUG-1997; 97WO-US014585.  
XX  
PR 19-AUG-1996; 96US-00699591.  
PR 19-NOV-1996; 96US-00753007.  
XX  
PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.  
XX  
PI Gearing DP, Busfield SJ;  
XX  
DR WPI; 1998-169084/15.  
DR P-PSDB; AAW48382.  
XX  
PT Mouse and human don-1 polypeptide(s) - useful for treatment of melanomas  
PT and adenocarcinoma(s), and for wound healing.  
XX  
PS Claim 4; Fig 4; 121pp; English.  
XX  
CC The sequence is that of a human don-1 gene splice variant. Don-1  
CC polypeptides stimulate proliferation of epithelial cells and thus are  
CC implicated in melanomas and adenocarcinomas in which epithelial cells  
CC proliferate out of control. Compounds that interfere with don-1 mediated  
CC cell proliferation can be used in the treatment of tumours such as  
CC melanomas and adenocarcinomas of the skin, oesophagus, lung, breast,  
CC liver, pancreas, gastrointestinal tract, colon, prostate or uterus.  
CC Alternatively, don-1 polypeptides can be used to stimulate epithelial  
CC cell proliferation, e.g. for wound healing  
XX  
SQ Sequence 1476 BP; 335 A; 475 C; 450 G; 216 T; 0 U; 0 Other;

Query Match 47.5%; Score 426.2; DB 2; Length 1476;

Best Local Similarity 89.6%; Pred. No. 8.6e-88;

Matches 458; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422  
|| || | || | | | | || | || | || | ||  
Db 98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 157  
Qy 423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482  
|||||  
Db 158 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217  
Qy 483 ATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542

Db	218		ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA	277
Qy	543		TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA	602
Db	278		TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA	337
Qy	603		CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC	662
Db	338		CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC	397
Qy	663		CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC	722
Db	398		CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC	457
Qy	723		CACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG	782
Db	458		CACCCTGTCATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG	517
Qy	783		CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC	842
Db	518		CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC	577
Qy	843		TGTGGGATACACCGGGGACAGGTGTCAGCAG	873
Db	578		AAATGGATTCTTCGCACAGAGATGTTTGGAG	608

RESULT 12

ABL40993

ID ABL40993 standard; DNA; 1054 BP.

XX

AC ABL40993;

XX

DT 03-JUL-2002 (first entry)

XX

DE Human neuregulin 2 gene exon 1.

XX

KW Human; MUC1; mucin; glycoprotein; cytostatic; cancer; tumour; ECD;

KW extracellular domain; neuregulin 2; isoform; gene; ds.

XX

OS Homo sapiens.

XX

PN WO200222685-A2.

XX

PD 21-MAR-2002.

XX

PF 11-SEP-2001; 2001WO-US028548.

XX

PR 11-SEP-2000; 2000US-0231841P.

XX

PA (KUFE/) KUFE D W.

PA (OHNO/) OHNO T.

XX

PI Kufe DW, Ohno T;

XX

DR WPI; 2002-339864/37.

XX  
PT Use of a mucin glycoprotein (MUC1) extracellular domain antagonist for  
PT manufacturing a medicant that inhibits the proliferation of MUC-1  
PT expressing cancer cells and that can treat cancers and reduce tumor  
PT growth.  
XX  
PS Disclosure; Page 61-62; 74pp; English.  
XX  
CC The invention relates to the use of a MUC1 (mucin glycoprotein)  
CC extracellular domain (ECD) antagonist for the manufacture of a medicant  
CC to inhibit the proliferation of MUC-1 expressing cancer cells. MUC1 ECD  
CC antagonists (optionally combined with a pharmaceutical carrier) can be  
CC administered to inhibit proliferation of MUC1-expressing cancer cells,  
CC useful to treat cancers e.g. skin cancer, prostate cancer and leukemia,  
CC especially in humans. The method may also be combined with administration  
CC of a chemotherapeutic agent (e.g. an alkylating agent, topoisomerase etc)  
CC or radiation to treat cancer, especially to reduce tumour growth. The  
CC polypeptides are also useful in screening to identify MUC1 ECD  
CC antagonists. The present sequence represents an exon fragment of the  
CC human neuregulin 2 gene  
XX  
SQ Sequence 1054 BP; 178 A; 367 C; 361 G; 148 T; 0 U; 0 Other;

Query Match 47.3%; Score 424; DB 6; Length 1054;  
Best Local Similarity 100.0%; Pred. No. 2.5e-87;  
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC	60
Db	589	ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC	648
Qy	61	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	120
Db	649	TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG	708
Qy	121	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	180
Db	709	GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	768
Qy	181	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	769	CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	828
Qy	241	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCG	300
Db	829	GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCG	888
Qy	301	CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	889	CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	948
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	949	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	1008
Qy	421	TGCG	424



Db 1009 TGCG 1012

RESULT 13

AAV17813

ID AAV17813 standard; cDNA; 1607 BP.

XX

AC AAV17813;

XX

DT 17-AUG-1998 (first entry)

XX

DE Mus musculus don-1 gene splice variant.

XX

KW Murine; don-1 gene; melanoma; treatment; adenocarcinoma; epithelial cell;  
KW proliferation; stimulation; treatment; tumours; skin; oesophagus; lung;  
KW breast; liver; pancreas; colon; prostate; gastrointestinal tract; uterus;  
KW wound healing; secreted protein; ss.

XX

OS Mus musculus.

XX

FH Key Location/Qualifiers

FT CDS 79. .624

FT /\*tag= a

FT /note= "secreted don-1 polypeptide"

XX

PN WO9807736-A1.

XX

PD 26-FEB-1998.

XX

PF 18-AUG-1997; 97WO-US014585.

XX

PR 19-AUG-1996; 96US-00699591.

PR 19-NOV-1996; 96US-00753007.

XX

PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.

XX

PI Gearing DP, Busfield SJ;

XX

DR WPI; 1998-169084/15.

DR P-PSDB; AAW48380.

XX

PT Mouse and human don-1 polypeptide(s) - useful for treatment of melanomas  
PT and adenocarcinoma(s), and for wound healing.

XX

PS Claim 4; Fig 2; 121pp; English.

XX

CC The sequence is that of a murine don-1 gene splice variant. Don-1  
CC polypeptides stimulate proliferation of epithelial cells and thus are  
CC implicated in melanomas and adenocarcinomas in which epithelial cells  
CC proliferate out of control. Compounds that interfere with don-1 mediated  
CC cell proliferation can be used in the treatment of tumours such as  
CC melanomas and adenocarcinomas of the skin, oesophagus, lung, breast,  
CC liver, pancreas, gastrointestinal tract, colon, prostate or uterus.  
CC Alternatively, don-1 polypeptides can be used to stimulate epithelial  
CC cell proliferation, e.g. for wound healing

XX

SQ Sequence 1607 BP; 365 A; 500 C; 480 G; 262 T; 0 U; 0 Other;

Query Match 45.2%; Score 405.4; DB 2; Length 1607;  
Best Local Similarity 87.9%; Pred. No. 5e-83;  
Matches 442; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

[illegible]

RESULT 14

AAV17812

ID     AAV17812 standard; cDNA; 2467 BP.

XX

AC AAV17812;

XX

DT 17-AUG-1998 (first entry)

XX

DE Mus musculus don-1 gene splice variant.

XX

KW Murine; don-1 gene; melanoma; treatment; adenocarcinoma; epithelial cell;

KW proliferation; stimulation; treatment; tumours; skin; oesophagus; lung;

KW breast; liver; pancreas; colon; prostate; gastrointestinal tract; uterus;

KW wound healing; transmembrane; ss.

XX

```

OS      Mus musculus.
XX
FH      Key                Location/Qualifiers
FT      CDS                79..1896
FT                        /*tag= a
FT                        /note= "transmembrane don-1 polypeptide"
XX
PN      WO9807736-A1.
XX
PD      26-FEB-1998.
XX
PF      18-AUG-1997;      97WO-US014585.
XX
PR      19-AUG-1996;      96US-00699591.
PR      19-NOV-1996;      96US-00753007.
XX
PA      (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX
PI      Gearing DP,      Busfield SJ;
XX
DR      WPI; 1998-169084/15.
DR      P-PSDB; AAW48379.
XX
PT      Mouse and human don-1 polypeptide(s) - useful for treatment of melanomas
PT      and adenocarcinoma(s), and for wound healing.
XX
PS      Claim 4; Fig 1; 121pp; English.
XX
CC      The sequence is that of a murine don-1 gene splice variant. Don-1
CC      polypeptides stimulate proliferation of epithelial cells and thus are
CC      implicated in melanomas and adenocarcinomas in which epithelial cells
CC      proliferate out of control. Compounds that interfere with don-1 mediated
CC      cell proliferation can be used in the treatment of tumours such as
CC      melanomas and adenocarcinomas of the skin, oesophagus, lung, breast,
CC      liver, pancreas, gastrointestinal tract, colon, prostate or uterus.
CC      Alternatively, don-1 polypeptides can be used to stimulate epithelial
CC      cell proliferation, e.g. for wound healing
XX
SQ      Sequence 2467 BP; 592 A; 752 C; 706 G; 417 T; 0 U; 0 Other;

Query Match          44.8%;  Score 402.2;  DB 2;  Length 2467;
Best Local Similarity 87.5%;  Pred. No. 3e-82;
Matches 440;  Conservative 0;  Mismatches 63;  Indels 0;  Gaps 0;

QY      371 CCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 430
      | |||||
DB      2 CTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 61

QY      431 GGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGA 490
      |||||
DB      62 GGCCCAAGCTGAAGAAGATGAAGAGCCAGACAGGAGAGGTGGGTGAGAAGCAGTCGCTCA 121

QY      491 AGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCTACCGTTGGTTCAAGGATGGCAAGG 550
      |||||
DB      122 AGTGTGAGGCAGCGCGGGAAACCCCCAGCCCTCCTATCGCTGGTTCAAGGATGGCAAGG 181

QY      551 AGCTCAACCGCAGCCGAGACATTGCGATCAAATATGGCAACGGCAGAAAGAACTCACGAC 610

```

Db	182	AACTCAACCGGAGTCGTGATATTTCGCATCAAGTATGGCAATGTCAGAAAGAACTCACGGC	241
Qy	611	TACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACA	670
Db	242	TACAGTTCAACAAAGTGAGCGTGGAGGATGCCGGGGAGTACGTCTGTGAGGCCGAGAACA	301
Qy	671	TCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGT	730
Db	302	TCCTTGGGAAGGACACCGTGAGGGGCCGACTCCATGTCAACAGCGTGACCACCCTCTGT	361
Qy	731	CATCCTGGTCTGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATG	790
Db	362	CATCCTGGTCTGGGACATGCCCCGGAAGTGCAATGAGACCGCCAAGTCCTACTGTGTGAATG	421
Qy	791	GAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAAGTGTCCTGTGGGAT	850
Db	422	GAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAACGGAT	481
Qy	851	ACACCGGGGACAGGTGTCAGCAG	873
Db	482	TCTTCGGACAGAGATGTTTGGAG	504



Db	301	TCCTTGGGAAGGACACCGTGAGGGGCCGACTCCATGTCAACAGCGTGAGCACCCTCTGT	360
Qy	731	CATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATG	790
Db	361	CATCCTGGTCGGGACATGCCCGGAAGTGCAATGAGACCGCCAAGTCCTACTGTGTGAATG	420
Qy	791	GAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCTGTGGGAT	850
Db	421	GAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAACGGAT	480
Qy	851	ACACCGGGGACAGGTGTCAGCAG	873
Db	481	TCTTCGGACAGAGATGTTTGGAG	503

Search completed: August 15, 2004, 05:47:25  
Job time : 433.237 secs

OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 05:20:54 ; Search time 82.0629 Seconds  
(without alignments)  
6065.966 Million cell updates/sec

Title: US-09-864-675-3  
Perfect score: 897  
Sequence: 1 atgaggcgcgacccggcccc.....caatggtcaacttctcctaa 897

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	835.4	93.1	1884	3	US-08-753-007A-5	Sequence 5, Appli
2	835.4	93.1	1884	3	US-09-398-496-5	Sequence 5, Appli
3	784	87.4	3441	2	US-08-525-864A-1	Sequence 1, Appli
4	738.6	82.3	993	2	US-08-525-864A-3	Sequence 3, Appli
5	427.8	47.7	1476	3	US-08-753-007A-7	Sequence 7, Appli
6	427.8	47.7	1476	3	US-09-398-496-7	Sequence 7, Appli
7	427.8	47.7	2268	3	US-08-753-007A-31	Sequence 31, Appl
8	427.8	47.7	2268	3	US-09-398-496-31	Sequence 31, Appl
9	405.4	45.2	1607	3	US-08-753-007A-3	Sequence 3, Appli
10	405.4	45.2	1607	3	US-09-398-496-3	Sequence 3, Appli
11	405.4	45.2	2467	3	US-08-753-007A-1	Sequence 1, Appli

12	405.4	45.2	2467	3	US-09-398-496-1	Sequence 1, Appli
13	216.2	24.1	1207	2	US-08-525-864A-5	Sequence 5, Appli
14	84	9.4	2003	1	US-08-036-555B-21	Sequence 21, Appl
15	84	9.4	2003	1	US-08-469-569-21	Sequence 21, Appl
16	84	9.4	2003	1	US-08-249-322A-21	Sequence 21, Appl
17	84	9.4	2003	1	US-08-469-526A-21	Sequence 21, Appl
18	84	9.4	2003	2	US-08-734-591A-21	Sequence 21, Appl
19	84	9.4	2003	2	US-08-469-660-21	Sequence 21, Appl
20	84	9.4	2003	3	US-08-341-018-71	Sequence 71, Appl
21	84	9.4	2003	3	US-08-470-335-21	Sequence 21, Appl
22	84	9.4	2003	3	US-08-735-021-21	Sequence 21, Appl
23	84	9.4	2003	3	US-08-734-664A-21	Sequence 21, Appl
24	84	9.4	2003	3	US-08-470-339-21	Sequence 21, Appl
25	84	9.4	2003	4	US-08-467-602-21	Sequence 21, Appl
26	84	9.4	2003	5	PCT-US94-05083C-21	Sequence 21, Appl
27	84	9.4	2003	5	PCT-US95-06846A-21	Sequence 21, Appl
28	83.4	9.3	1108	1	US-08-036-555B-135	Sequence 135, App
29	83.4	9.3	1108	1	US-08-469-569-135	Sequence 135, App
30	83.4	9.3	1108	1	US-08-249-322A-135	Sequence 135, App
31	83.4	9.3	1108	1	US-08-469-526A-135	Sequence 135, App
32	83.4	9.3	1108	2	US-08-734-591A-135	Sequence 135, App
33	83.4	9.3	1108	2	US-08-469-660-135	Sequence 135, App
34	83.4	9.3	1108	3	US-08-341-018-5	Sequence 5, Appli
35	83.4	9.3	1108	3	US-08-470-335-135	Sequence 135, App
36	83.4	9.3	1108	3	US-08-735-021-135	Sequence 135, App
37	83.4	9.3	1108	3	US-08-734-664A-135	Sequence 135, App
38	83.4	9.3	1108	3	US-08-470-339-135	Sequence 135, App
39	83.4	9.3	1108	4	US-08-467-602-135	Sequence 135, App
40	83.4	9.3	1108	5	PCT-US94-05083C-131	Sequence 131, App
41	83.4	9.3	1108	5	PCT-US95-06846A-135	Sequence 135, App
42	78.4	8.7	1193	1	US-08-469-526A-134	Sequence 134, App
43	78.4	8.7	1193	2	US-08-734-591A-134	Sequence 134, App
44	78.4	8.7	1193	3	US-08-341-018-3	Sequence 3, Appli
45	78.4	8.7	1193	3	US-08-470-335-134	Sequence 134, App

# ALIGNMENTS

## RESULT 1

US-08-753-007A-5

; Sequence 5, Application US/08753007A

; Patent No. 6074841

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; TITLE OF INVENTION: AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:



```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/753,007A
; FILING DATE: 19-NOV-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1884 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 664...1883
; OTHER INFORMATION:
US-08-753-007A-5

```

```

Query Match          93.1%; Score 835.4; DB 3; Length 1884;
Best Local Similarity 98.1%; Pred. No. 2.3e-201;
Matches 856; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

```

```

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 60
          |||
Db      218 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 277

Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
          |||
Db      278 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 337

Qy      121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
          |||
Db      338 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 397

Qy      181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
          |||
Db      398 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 457

Qy      241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
          |||
Db      458 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 517

Qy      301 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360

```

Db	518	CGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	577
Qy	361	CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	578	CCCCT-GATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC	636
Qy	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	637	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	696
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	697	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	756
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAAG	600
Db	757	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAAG	816
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	817	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	876
Qy	661	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	877	GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	936
Qy	721	ACCACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	780
Db	937	ACCACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT	996
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT	840
Db	997	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT	1056
Qy	841	CCTGTGGGATACACCGGGGACAGGTGTCAGCAG	873
Db	1057	CCAAATGGATTCTTCGGACAGAGATGTTTGGAG	1089

RESULT 2

US-09-398-496-5

; Sequence 5, Application US/09398496

; Patent No. 6133423

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; TITLE OF INVENTION: AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804



Db 458 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 517  
 Qy 301 CGCTACATCTTTTTCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 518 CGCTACATCTTTTTCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 577  
 Qy 361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420  
 ||||| |||||||||||||||||||||||||||||||||||||||||||||  
 Db 578 CCCCT-GATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC 636  
 Qy 421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 637 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 696  
 Qy 481 CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 697 CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 756  
 Qy 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAAG 600  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 757 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAAG 816  
 Qy 601 AACTCAGACTACAGTTCAACAAGGTGAAGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 817 AACTCAGACTACAGTTCAACAAGGTGAAGTGGAGGACGCTGGGGAGTATGTCTGCGAG 876  
 Qy 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 877 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 936  
 Qy 721 ACCACCCTGTCATCCTGGTCGGGGCACGCCCCGAAGTGCAACGAGACAGCCAAGTCCTAT 780  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 937 ACCACCCTGTCATCCTGGTCGGGGCACGCCCCGAAGTGCAACGAGACAGCCAAGTCCTAT 996  
 Qy 781 TCGCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 997 TCGCTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 1056  
 Qy 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873  
 || ||||| | ||| | || ||| ||  
 Db 1057 CCAAATGGATTCTTCGGACAGAGATGTTTGGAG 1089

RESULT 3

US-08-525-864A-1

; Sequence 1, Application US/08525864A

; Patent No. 5912326

; GENERAL INFORMATION:

; APPLICANT: Chang, Han

; TITLE OF INVENTION: Cerebellum-derived Growth Factors, and Uses

; TITLE OF INVENTION: Related thereto

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 28 State Street

; CITY: Boston

; STATE: Massachusetts

```

; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: AscII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,864A
; FILING DATE: 8-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kara, Catherine J.
; REGISTRATION NUMBER: 41,106
; REFERENCE/DOCKET NUMBER: HUI-017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3441 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 180..2441
US-08-525-864A-1

```

```

Query Match          87.4%; Score 784; DB 2; Length 3441;
Best Local Similarity 92.2%; Pred. No. 2.5e-188;
Matches 826; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

```

```

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 60
        |||
Db      180 ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 239
        |||

Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
        |||
Db      240 TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 299
        |||

Qy      121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
        |||
Db      300 GGCAAGGTACAGGGACTGGCCCCGGCAGGCGGTTCCAGCTCTAACAGCACCCGAGAGCCT 359
        |||

Qy      181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
        |||
Db      360 CCCGCCTCGGGTCGGGTGGCGCTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 419
        |||

Qy      241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
        |||
Db      420 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG 479
        |||

Qy      301 CGCTACATCTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
        |||
Db      480 CGCTACATCTTTTTCCTGGAGCCACCGAGCAGCCCTTAGTTTTTAAGACAGCCTTTGCC 539
        |||

```



```

;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  AscII (text)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/525,864A
;   FILING DATE:  8-SEP-1995
;   CLASSIFICATION:  530
;   ATTORNEY/AGENT INFORMATION:
;   NAME:  Kara, Catherine J.
;   REGISTRATION NUMBER:  41,106
;   REFERENCE/DOCKET NUMBER:  HUI-017
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (617)227-7400
;   TELEFAX:  (617)742-4214
;   INFORMATION FOR SEQ ID NO:  3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:  993 base pairs
;   TYPE:  nucleic acid
;   STRANDEDNESS:  double
;   TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA
;   FEATURE:
;   NAME/KEY:  CDS
;   LOCATION:  1..990
US-08-525-864A-3

```

```

Query Match          82.3%;  Score 738.6;  DB 2;  Length 993;
Best Local Similarity  90.4%;  Pred. No. 4.7e-177;
Matches 789;  Conservative  0;  Mismatches  84;  Indels  0;  Gaps  0;

```

```

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCTGCTCGCCTGC 60
        |||
Db      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCGATGCTGCTCTTCGGTGTGTCACTCGCCTGC 60

Qy     61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
        |||
Db     61 TACTCGCCCAGCCTCAAGTCCGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120

Qy    121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
        |||
Db    121 GGCAAGGTACAGGGAAGTGGCCCCGGCAGGCGGTTCAGCTCTAACAGCACCCGAGAGCCT 180

Qy    181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
        |||
Db    181 CCCGCCTCGGGTCGGGTGGCGTGGTGAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240

Qy    241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
        |||
Db    241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGCGCGCCGCTCGAAAGGAACCAG 300

Qy    301 CGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
        |||
Db    301 CGCTACATCTTTTTCCTGGAGCCCACCGAGCAGCCCTTAGTCTTTAAGACAGCCTTTGCC 360

Qy    361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
        ||
Db    361 CCGGTCGACCCTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420

```

Qy	421	TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG	480
Db	421	TGCGCAACCCGGCCCAAGCTGAAGAAGATGAAGAGTCAGACAGGAGAGGTGGGCGAGAAG	480
Qy	481	CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG	540
Db	481	CAGTCGCTCAAGTGTGAGGCGGCGGCGGGGAACCCCCAGCCCTCCTATCGATGGTTCAAG	540
Qy	541	GATGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAG	600
Db	541	GACGGCAAGGAGCTCAACCGGAGTCGTGACATTTCGCATCAAGTATGGCAACGGCAGAAAG	600
Qy	601	AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG	660
Db	601	AACTCACGGCTACAGTTCAACAAAGTGAAGGTGGAGGACGCTGGAGAGTACGTCTGTGAG	660
Qy	661	GCCGAGAACATCCTTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC	720
Db	661	GCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGTGAGC	720
Qy	721	ACCACCCTGTGCATCCTGGTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTAT	780
Db	721	ACCACTCTGTTCGTCTGGTCGGGGCACGCCCAGGAGTGAATGAGACAGCCAAGTCCTAC	780
Qy	781	TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT	840
Db	781	TGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAAAATGT	840
Qy	841	CCTGTGGGATACACCGGGGACAGGTGTCAGCAG	873
Db	841	CCAAACGGATTCTTCGGACAGAGATGTTTGGAG	873

## RESULT 5

US-08-753-007A-7

; Sequence 7, Application US/08753007A

; Patent No. 6074841

## ; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; TITLE OF INVENTION: AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

## ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

## ; COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Diskette
```

```
;      COMPUTER:  IBM Compatible
```

```
; OPERATING SYSTEM:  DOS
```

```
; SOFTWARE: FastSEQ Version 2.0
```

```
; CURRENT APPLICATION DATA:
```

; APPLICATION NUMBER: US/08/753,007A



```

; FILING DATE: 19-NOV-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1476 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 69...1475
; OTHER INFORMATION:
US-08-753-007A-7

```

```

Query Match          47.7%; Score 427.8; DB 3; Length 1476;
Best Local Similarity 89.8%; Pred. No. 1.3e-98;
Matches 459; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```

```

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
      || || | || | | | | || | | | | | | | |
Db      98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
      |||||
Db      158 AGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
      |||||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 602
      |||||
Db      278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 337

Qy      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
      |||||
Db      338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

Qy      663 CGAGAACATCCTGGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
      |||||
Db      398 CGAGAACATCCTGGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

Qy      723 CACCCTGTCATCCTGGTTCGGGGCACGCCCCGAAGTGCAACGAGACAGCCAAGTCCTATTG 782
      |||||
Db      458 CACCCTGTCATCCTGGTTCGGGGCACGCCCCGAAGTGCAACGAGACAGCCAAGTCCTATTG 517

```

QY 783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842  
 |||  
 Db 518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577  
 QY 843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873  
 |||  
 Db 578 AAATGGATTCTTCGGACAGAGATGTTTGGAG 608

RESULT 6

US-09-398-496-7

; Sequence 7, Application US/09398496

; Patent No. 6133423

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; TITLE OF INVENTION: AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/398,496

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/753,007

; FILING DATE: 19-NOV-1996

; APPLICATION NUMBER: 08/699,591

; FILING DATE: 19-AUG-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Fasse, J. Peter

; REGISTRATION NUMBER: 32,983

; REFERENCE/DOCKET NUMBER: 07334/022001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-542-5070

; TELEFAX: 617-542-8906

; TELEX:

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1476 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

```

;      NAME/KEY:  Coding Sequence
;      LOCATION:  69...1475
;      OTHER INFORMATION:
US-09-398-496-7

```

## RESULT 7

```

;   ADDRESSEE:  Fish & Richardson P.C.
;   STREET:    225 Franklin Street
;   CITY:     Boston
;   STATE:    MA
;   COUNTRY:   US
;   ZIP:      02110-2804
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Diskette
;   COMPUTER:   IBM Compatible
;   OPERATING SYSTEM:  DOS
;   SOFTWARE:   FastSEQ Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/753,007A
;   FILING DATE:   19-NOV-1996
;   CLASSIFICATION:  536
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  08/699,591
;   FILING DATE:   19-AUG-1996
;   ATTORNEY/AGENT INFORMATION:
;   NAME:        Fasse, J. Peter
;   REGISTRATION NUMBER:  32,983
;   REFERENCE/DOCKET NUMBER:  07334/022001
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:   617-542-5070
;   TELEFAX:    617-542-8906
;   TELEX:
;   INFORMATION FOR SEQ ID NO:  31:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:     2268 base pairs
;   TYPE:       nucleic acid
;   STRANDEDNESS:  single
;   TOPOLOGY:   linear
;   MOLECULE TYPE:  cDNA
;   FEATURE:
;   NAME/KEY:   Coding Sequence
;   LOCATION:   69...2009
;   OTHER INFORMATION:
US-08-753-007A-31

```

```

Query Match          47.7%;  Score 427.8;  DB 3;  Length 2268;
Best Local Similarity 89.8%;  Pred. No. 1.5e-98;
Matches 459;  Conservative 0;  Mismatches 52;  Indels 0;  Gaps 0;

```

```

Qy      363 CCTCGATACCAACGGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
      || || | || | | | | | | | | | | | | | | | |
Db      98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 157

Qy      423 CGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
      |||||
Db      158 AGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

Qy      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
      |||||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

Qy      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTGCGCATCAAATATGGCAACGGCAGAAAGAA 602
      |||||

```

Db 278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 337  
 Qy 603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662  
 |||||  
 Db 338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397  
 Qy 663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722  
 |||||  
 Db 398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457  
 Qy 723 CACCCTGTCATCCTGGTGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCATTG 782  
 |||||  
 Db 458 CACCCTGTCATCCTGGTGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCATTG 517  
 Qy 783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842  
 |||||  
 Db 518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577  
 Qy 843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873  
 |||| | || | || || ||  
 Db 578 AAATGGATTCTTCGGACAGAGATGTTTGGAG 608

RESULT 8

US-09-398-496-31

; Sequence 31, Application US/09398496

; Patent No. 6133423

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; TITLE OF INVENTION: AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/398,496

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/753,007

; FILING DATE: 19-NOV-1996

; APPLICATION NUMBER: 08/699,591

; FILING DATE: 19-AUG-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Fasse, J. Peter

; REGISTRATION NUMBER: 32,983

```

; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX:
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2268 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 69...2009
; OTHER INFORMATION:
US-09-398-496-31

```

```

Query Match          47.7%; Score 427.8; DB 3; Length 2268;
Best Local Similarity 89.8%; Pred. No. 1.5e-98;
Matches 459; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```

```

QY      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
      || || | || | | | | | | | | | | | | | | | |
Db      98 CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCGATCCCGGGGAGAA 157

QY      423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482
      |||||
Db      158 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217

QY      483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542
      |||||
Db      218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277

QY      543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 602
      |||||
Db      278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 337

QY      603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662
      |||||
Db      338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397

QY      663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722
      |||||
Db      398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457

QY      723 CACCCTGTCATCCTGGTTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTATTG 782
      |||||
Db      458 CACCCTGTCATCCTGGTTCGGGGCACGCCCAGGAGTGAACGAGACAGCCAAGTCCTATTG 517

QY      783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842
      |||||
Db      518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAATGTCC 577

QY      843 TGTGGGATACACCGGGGACAGGTGTGAGCAG 873
      |||| | || | || | || |
Db      578 AAATGGATTCTTCGGACAGAGATGTTTGGAG 608

```

RESULT 9  
 US-08-753-007A-3  
 ; Sequence 3, Application US/08753007A  
 ; Patent No. 6074841  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gearing, David P.  
 ; APPLICANT: Busfield, Samantha J.  
 ; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES  
 ; TITLE OF INVENTION: AND USES THEREFOR  
 ; NUMBER OF SEQUENCES: 33  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson P.C.  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: US  
 ; ZIP: 02110-2804  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSEQ Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/753,007A  
 ; FILING DATE: 19-NOV-1996  
 ; CLASSIFICATION: 536  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/699,591  
 ; FILING DATE: 19-AUG-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Fasse, J. Peter  
 ; REGISTRATION NUMBER: 32,983  
 ; REFERENCE/DOCKET NUMBER: 07334/022001  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617-542-5070  
 ; TELEFAX: 617-542-8906  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1607 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: Coding Sequence  
 ; LOCATION: 79...621  
 ; OTHER INFORMATION:  
 US-08-753-007A-3

Query Match 45.2%; Score 405.4; DB 3; Length 1607;  
 Best Local Similarity 87.9%; Pred. No. 5.8e-93;  
 Matches 442; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 371 CCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 430

```

      | |||||
Db      2 CTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 61

Qy      431 GGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGA 490
      | |||||
Db      62 GGCCCAAGCTGAAGAAGATGAAGAGCCAGACAGGAGAGGTGGGTGAGAAGCAGTCGCTCA 121

Qy      491 AGTGTGAGGCAGCAGCCGGAATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGG 550
      | |||||
Db      122 AGTGTGAGGCAGCGGCGGGAAACCCCCAGCCCTCCTATCGCTGGTTCAAGGATGGCAAGG 181

Qy      551 AGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGAC 610
      | |||||
Db      182 AACTCAACCGGAGTCGTGATATTTCGCATCAAGTATGGCAATGTCAGAAAGAACTCACGGC 241

Qy      611 TACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACA 670
      | |||||
Db      242 TACAGTTCAACAAAGTGAGGGTGGAGGATGCCGGGGAGTACGTCTGTGAGGCCGAGAACA 301

Qy      671 TCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGT 730
      | |||||
Db      302 TCCTTGGGAAGGACACCGTGAGGGGCCGACTCCATGTCAACAGCGTGAGCACCACCTCTGT 361

Qy      731 CATCCTGGTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATG 790
      | |||||
Db      362 CATCCTGGTCGGGACATGCCCGGAAGTGCAATGAGACCGCCAAGTCCTACTGTGTGAATG 421

Qy      791 GAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCTGTGGGAT 850
      | |||||
Db      422 GAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAAATGTCCAAACGGAT 481

Qy      851 ACACCGGGGACAGGTGTCAGCAG 873
      | |||
Db      482 TCTTCGGACAGAGATGTTTGGAG 504

```

RESULT 10

US-09-398-496-3

; Sequence 3, Application US/09398496

; Patent No. 6133423

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; TITLE OF INVENTION: AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS







US-08-753-007A-1

Matches 442; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 851 ACACCGGGGACAGGTGTCAGCAG 873

US-09-398-496-1

; Patent No. 6133423

APPLICANT: Gearing, David P.

10 ; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

```

; TITLE OF INVENTION:  AND USES THEREFOR
; NUMBER OF SEQUENCES:  33
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  Fish & Richardson P.C.
;   STREET:  225 Franklin Street
;   CITY:  Boston
;   STATE:  MA
;   COUNTRY:  US
;   ZIP:  02110-2804
; COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Diskette
;   COMPUTER:  IBM Compatible
;   OPERATING SYSTEM:  DOS
;   SOFTWARE:  FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/09/398,496
;   FILING DATE:
;   CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  08/753,007
;   FILING DATE:  19-NOV-1996
;   APPLICATION NUMBER:  08/699,591
;   FILING DATE:  19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
;   NAME:  Fasse, J. Peter
;   REGISTRATION NUMBER:  32,983
;   REFERENCE/DOCKET NUMBER:  07334/022001
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  617-542-5070
;   TELEFAX:  617-542-8906
;   TELEX:
; INFORMATION FOR SEQ ID NO:  1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH:  2467 base pairs
;   TYPE:  nucleic acid
;   STRANDEDNESS:  single
;   TOPOLOGY:  circular
; MOLECULE TYPE:  cDNA
; FEATURE:
;   NAME/KEY:  Coding Sequence
;   LOCATION:  79...1893
;   OTHER INFORMATION:
US-09-398-496-1

```

```

Query Match          45.2%;  Score 405.4;  DB 3;  Length 2467;
Best Local Similarity 87.9%;  Pred. No. 6.7e-93;
Matches 442;  Conservative 0;  Mismatches 61;  Indels 0;  Gaps 0;

```

```

QY      371 CCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 430
          | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2   CTAACGGCAAAAACATCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTGCGCCACCC 61

QY      431 GGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAATCGCTGA 490
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      62  GGCCCAAGCTGAAGAAGATGAAGAGCCAGACAGGAGAGGTGGGTGAGAAGCAGTCGCTCA 121

QY      491 AGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGATGGCAAGG 550

```

Db	122	AGTGTGAGGCAGCGCGGGAAACCCCGAGCCCTCCTATCGCTGGTTCAAGGATGGCAAGG	181
Qy	551	AGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAACTCACGAC	610
Db	182	AACTCAACCGGAGTCGTGATATTTCGCATCAAGTATGGCAATGTCAGAAAGAACTCACGGC	241
Qy	611	TACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGCCGAGAACA	670
Db	242	TACAGTTCAACAAAGTGAGGGTGGAGGATGCCGGGGAGTACGTCTGTGAGGCCGAGAACA	301
Qy	671	TCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCACCACCCTGT	730
Db	302	TCCTTGGGAAGGACACCGTGAGGGGCCGACTCCATGTCAACAGCGTGAGCACCCTCTGT	361
Qy	731	CATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTGCGTCAATG	790
Db	362	CATCCTGGTCGGGACATGCCCGGAAGTGCAATGAGACCGCCAAGTCCTACTGTGTGAATG	421
Qy	791	GAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCCTGTGGGAT	850
Db	422	GAGGCGTGTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCCAAACGGAT	481
Qy	851	ACACCGGGGACAGGTGTCAGCAG	873
Db	482	TCTTCGACAGAGATGTTTGGAG	504

RESULT 13

US-08-525-864A-5

; Sequence 5, Application US/08525864A

; Patent No. 5912326

; GENERAL INFORMATION:

; APPLICANT: Chang, Han

; TITLE OF INVENTION: Cerebellum-derived Growth Factors, and Uses

; TITLE OF INVENTION: Related thereto

; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 28 State Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: AscII (text)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/525,864A

; FILING DATE: 8-SEP-1995

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Kara, Catherine J.

; REGISTRATION NUMBER: 41,106

; REFERENCE/DOCKET NUMBER: HUI-017

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..394
US-08-525-864A-5

```

Query Match 24.1%; Score 216.2; DB 2; Length 1207;  
Best Local Similarity 86.3%; Pred. No. 2.7e-45;  
Matches 239; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy	597	AAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTG	656
Db	1	AAAGAACTCACGGCTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGAGAGTACGTCTG	60
Qy	657	CGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGT	716
Db	61	TGAGGCTGAGAACATCCTTGGGAAGGACACTGTGAGGGGCCGGCTCCATGTCAACAGTGT	120
Qy	717	GAGCACCACTCTGTCTCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTC	776
Db	121	GAGCACCACTCTGTCTCCTGGTTCGGGGCACGCCCCGGAAGTGCAATGAGACAGCCAAGTC	180
Qy	777	CTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAA	836
Db	181	CTACTGTGTGAATGGAGGCGTGTGCTACTACATCGAAGGCATCAACCAACTCTCCTGCAA	240
Qy	837	GTGTCCTGTGGGATACACCGGGGACAGGTGTCAGCAG	873
Db	241	ATGTCCAAACGGATTCTTCGGACAGAGATGTTTGGAG	277

RESULT 14

US-08-036-555B-21

```
; Sequence 21, Application US/08036555B
```

; Patent No. 5530109

; GENERAL INFORMATION:

; APPLICANT: Goodearl, Andrew; Stroobant, Paul;

; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;

; APPLICANT: Chen, Maio Su; Hiles, Ian

; TITLE OF INVENTION: Glial Mitogenic Factors, Their

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

```
; NUMBER OF SEQUENCES: 184
```

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Felfe &amp; Lynch

STREET: 805 Third Avenue

; CITY: New York City

; STATE: New York

; COUNTRY: USA

```

; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/036,555B
; FILING DATE: 24-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APRIL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LUD 5250.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2003
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; OTHER INFORMATION: N in positions 31 and 32 could be either
; OTHER INFORMATION: A or G.
US-08-036-555B-21

```

```

Query Match          9.4%; Score 84; DB 1; Length 2003;
Best Local Similarity 49.0%; Pred. No. 6.9e-12;
Matches 361; Conservative 0; Mismatches 345; Indels 30; Gaps 4;

```

```

QY      189 GGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGGGGGCTGCA 248
      || ||| || | | || || ||| ||| | | || | |||
Db      762 GGTGTGGGCGGTGAAAGCCGGGGGCTTGAAGAAGGACTCGCTGCTCACCGTGCGCCTGGG 821

QY      249 GCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAGCGCTACAT 308
      | | | | | | | | | | | | | | | | | | | | | |
Db      822 GACCTGGGGCCACCCCGCCTTCCCCTCCTGCGGGAGGCTCAAGGAGGACAGCAGGTACAT 881

QY      309 CTTTTTCCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACG-----GCCTT 356
      ||| ||| ||||| || | | | | | | | | | |||

```





```

;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;      COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/469,569
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  530
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/036,555
;      FILING DATE:  24-MAR-1993
;      APPLICATION NUMBER:  07/965,173
;      FILING DATE:  23-OCT-1992
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/940,389
;      FILING DATE:  03-SEP-1992
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/907,138
;      FILING DATE:  30-JUN-1992
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/863,703
;      FILING DATE:  03-APRIL-1992
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.K. 91 07566.3
;      FILING DATE:  10-APRIL-1991
;      ATTORNEY/AGENT INFORMATION:
;      NAME:  Tsai, Christine H.
;      REGISTRATION NUMBER:  34,266
;      REFERENCE/DOCKET NUMBER:  LUD 5250.4
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;      INFORMATION FOR SEQ ID NO: 21:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH:  2003
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;      FEATURE:
;      OTHER INFORMATION:  N in positions 31 and 32 could be either
;      OTHER INFORMATION:  A or G.
US-08-469-569-21

```

```

Query Match          9.4%;  Score 84;  DB 1;  Length 2003;
Best Local Similarity  49.0%;  Pred. No. 6.9e-12;
Matches 361;  Conservative  0;  Mismatches 345;  Indels  30;  Gaps  4;

```

```

Qy      189 GGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGGGGGCTGCA 248
      ||  ||| || |  ||  || | |||  |||| | | || | |||
Db      762 GGTGTGGGCGGTGAAAGCCGGGGGCTTGAAGAAGGACTCGCTGCTCACCGTGCGCCTGGG 821
Qy      249 GCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAGCGCTACAT 308

```



OM nucleic - nucleic search, using sw model

Run on: August 15, 2004, 08:02:19 ; Search time 483.839 Seconds  
(without alignments)  
9096.466 Million cell updates/sec

Title: US-09-864-675-3  
Perfect score: 897  
Sequence: 1 atgaggcgcgacccggccccc.....caatgggtcaacttctcctaa 897

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

8

Result Query

No.	Score	Match	Length	DB	ID	Description
1	897	100.0	897	9	US-09-864-675-3	Sequence 3, Appli
2	849	94.6	994	9	US-09-864-675-1	Sequence 1, Appli
3	835.4	93.1	1884	14	US-10-096-241-5	Sequence 5, Appli
4	427.8	47.7	1476	14	US-10-096-241-7	Sequence 7, Appli
5	427.8	47.7	2268	14	US-10-096-241-31	Sequence 31, Appl
6	424.8	47.4	22693	13	US-10-271-416-6	Sequence 6, Appli
7	424	47.3	1054	16	US-10-447-839A-10	Sequence 10, Appl
8	405.4	45.2	1607	14	US-10-096-241-3	Sequence 3, Appli
9	405.4	45.2	2467	14	US-10-096-241-1	Sequence 1, Appli
10	173	19.3	201	15	US-10-029-386-26613	Sequence 26613, A
11	173	19.3	419	16	US-10-447-839A-11	Sequence 11, Appl
12	173	19.3	573	15	US-10-029-386-12913	Sequence 12913, A
13	173	19.3	45450	13	US-10-271-416-7	Sequence 7, Appli
14	124.6	13.9	493	16	US-10-447-839A-12	Sequence 12, Appl
15	124	13.8	350	16	US-10-447-839A-13	Sequence 13, Appl
c 16	113.6	12.7	579	15	US-10-029-386-2532	Sequence 2532, Ap
c 17	111.8	12.5	171	15	US-10-029-386-16232	Sequence 16232, A
18	84	9.4	1986	10	US-09-373-658-71	Sequence 71, Appl
19	84	9.4	1986	11	US-09-989-687-71	Sequence 71, Appl
20	84	9.4	2003	8	US-08-736-019-21	Sequence 21, Appl
21	84	9.4	2003	10	US-09-366-886-71	Sequence 71, Appl
22	84	9.4	2003	10	US-09-373-658-72	Sequence 72, Appl
23	84	9.4	2003	11	US-09-989-687-72	Sequence 72, Appl
24	83.4	9.3	1108	8	US-08-736-019-135	Sequence 135, App
25	83.4	9.3	1108	10	US-09-366-886-5	Sequence 5, Appli
26	78.4	8.7	1193	8	US-08-736-019-134	Sequence 134, App
27	78.4	8.7	1193	10	US-09-366-886-3	Sequence 3, Appli
28	76.8	8.6	1140	8	US-08-736-019-149	Sequence 149, App
29	76.8	8.6	1140	10	US-09-366-886-55	Sequence 55, Appl
30	72.8	8.1	744	8	US-08-736-019-133	Sequence 133, App
31	72.8	8.1	744	10	US-09-366-886-1	Sequence 1, Appli
32	58.8	6.6	310	16	US-10-447-839A-15	Sequence 15, Appl
33	58.8	6.6	848	13	US-10-027-632-158930	Sequence 158930,
34	58.8	6.6	848	16	US-10-027-632-158930	Sequence 158930,
35	54.6	6.1	3272	12	US-10-152-319A-1922	Sequence 1922, Ap
36	52	5.8	1092	15	US-10-156-761-4747	Sequence 4747, Ap
37	52	5.8	9025608	15	US-10-156-761-1	Sequence 1, Appli
38	51	5.7	162	15	US-10-076-816-8	Sequence 8, Appli
39	51	5.7	162	15	US-10-077-381-8	Sequence 8, Appli
40	51	5.7	162	17	US-10-639-779-1	Sequence 1, Appli
41	50.6	5.6	3111	9	US-09-773-517-12	Sequence 12, Appl
42	50.6	5.6	3111	9	US-09-792-025-12	Sequence 12, Appl
43	50.6	5.6	3111	9	US-09-849-868-12	Sequence 12, Appl
44	50.6	5.6	3111	15	US-10-290-578-1	Sequence 1, Appli
45	50.6	5.6	3111	15	US-10-453-183-12	Sequence 12, Appl

#### ALIGNMENTS

RESULT 1

US-09-864-675-3

; Sequence 3, Application US/09864675

; Patent No. US20020081286A1

; GENERAL INFORMATION:

```

; APPLICANT: Marchionni, Mark
; TITLE OF INVENTION: NRG-2 NUCLEIC ACID MOLECULES,
; TITLE OF INVENTION: POLYPEPTIDES, AND DIAGNOSTIC AND THERAPEUTIC METHODS
; FILE REFERENCE: 04585/049002
; CURRENT APPLICATION NUMBER: US/09/864,675
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/206,495
; PRIOR FILING DATE: 2000-05-23
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 897
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-864-675-3

```

```

Query Match          100.0%; Score 897; DB 9; Length 897;
Best Local Similarity 100.0%; Pred. No. 2.2e-244;
Matches 897; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 60

Qy     61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120

Qy    121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180

Qy    181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240

Qy    241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300

Qy    301 CGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CGCTACATCTTTTTCCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360

Qy    361 CCCCTCGATACCAACGGCAAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCCCTCGATACCAACGGCAAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420

Qy    421 TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TGCGCCACCCGGCCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480

Qy    481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCTACCGTTGGTTCAAG 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CAATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCTACCGTTGGTTCAAG 540

Qy    541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG 600

```

```

      |||
Db      541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAG 600
Qy      601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
      |||
Db      601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660
Qy      661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
      |||
Db      661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720
Qy      721 ACCACCCTGTCATCCTGGTGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
      |||
Db      721 ACCACCCTGTCATCCTGGTGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780
Qy      781 TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840
      |||
Db      781 TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840
Qy      841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGGTCAACTTCTCCTAA 897
      |||
Db      841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAGTTCGCAATGGTCAACTTCTCCTAA 897

```

RESULT 2

US-09-864-675-1

; Sequence 1, Application US/09864675

; Patent No. US20020081286A1

; GENERAL INFORMATION:

; APPLICANT: Marchionni, Mark

; TITLE OF INVENTION: NRG-2 NUCLEIC ACID MOLECULES,

; TITLE OF INVENTION: POLYPEPTIDES, AND DIAGNOSTIC AND THERAPEUTIC METHODS

; FILE REFERENCE: 04585/049002

; CURRENT APPLICATION NUMBER: US/09/864,675

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/206,495

; PRIOR FILING DATE: 2000-05-23

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 994

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-864-675-1

Query Match 94.6%; Score 849; DB 9; Length 994;

Best Local Similarity 98.3%; Pred. No. 9.7e-231;

Matches 858; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 60
      |||
Db      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTCGCTCGCCTGC 60
Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
      |||
Db      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120

```



```

; GENERAL INFORMATION:
;   APPLICANT: Gearing, David P.
;               Busfield, Samantha J.
;   TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
;                       AND USES THEREFOR
;   NUMBER OF SEQUENCES: 33
;   CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Fish & Richardson P.C.
;       STREET: 225 Franklin Street
;       CITY: Boston
;       STATE: MA
;       COUNTRY: US
;       ZIP: 02110-2804
;   COMPUTER READABLE FORM:
;       MEDIUM TYPE: Diskette
;       COMPUTER: IBM Compatible
;       OPERATING SYSTEM: DOS
;       SOFTWARE: FastSEQ Version 2.0
;   CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/10/096,241
;       FILING DATE: 12-Mar-2002
;       CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 08/699,591
;       FILING DATE: 19-AUG-1996
;   ATTORNEY/AGENT INFORMATION:
;       NAME: Fasse, J. Peter
;       REGISTRATION NUMBER: 32,983
;       REFERENCE/DOCKET NUMBER: 07334/022001
;   TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 617-542-5070
;       TELEFAX: 617-542-8906
;       TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 5:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 1884 base pairs
;           TYPE: nucleic acid
;           STRANDEDNESS: single
;           TOPOLOGY: linear
;       MOLECULE TYPE: cDNA
;       FEATURE:
;           NAME/KEY: Coding Sequence
;           LOCATION: 664...1883
;       OTHER INFORMATION:
;       SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-096-241-5

```

```

Query Match          93.1%;  Score 835.4;  DB 14;  Length 1884;
Best Local Similarity 98.1%;  Pred. No. 8.2e-227;
Matches 856;  Conservative 0;  Mismatches 16;  Indels 1;  Gaps 1;

```

```

Qy      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 60
        |||
Db      218 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCGCTCGCCTGC 277
        |||

Qy      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
        |||

```



Db 278 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 337

Qy 121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180  
 |||

Db 338 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 397

Qy 181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240  
 |||

Db 398 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 457

Qy 241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300  
 |||

Db 458 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 517

Qy 301 CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360  
 |||

Db 518 CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 577

Qy 361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420  
 |||

Db 578 CCCCT-GATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGGC 636

Qy 421 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 480  
 |||

Db 637 TGCGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAG 696

Qy 481 CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 540  
 |||

Db 697 CAATCGCTGAAGTGTGAGGCAGCAGCCGTAATCCCCAGCCTTCCTACCGTTGGTTCAAG 756

Qy 541 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAAG 600  
 |||

Db 757 GATGGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAGAAAAG 816

Qy 601 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 660  
 |||

Db 817 AACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAG 876

Qy 661 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 720  
 |||

Db 877 GCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGC 936

Qy 721 ACCACCCTGTCATCCTGGTTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 780  
 |||

Db 937 ACCACCCTGTCATCCTGGTTCGGGGCAGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTAT 996

Qy 781 TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGT 840  
 |||

Db 997 TGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGT 1056

Qy 841 CCTGTGGGATACACCGGGGACAGGTGTCAGCAG 873  
 || ||| || ||| ||| |||

Db 1057 CCAAATGGATTCTTCGGACAGAGATGTTTGGAG 1089

```

; Sequence 7, Application US/10096241
; Publication No. US20020127594A1
;   GENERAL INFORMATION:
;     APPLICANT: Gearing, David P.
;               Busfield, Samantha J.
;     TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
;                       AND USES THEREFOR
;     NUMBER OF SEQUENCES: 33
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Fish & Richardson P.C.
;       STREET: 225 Franklin Street
;       CITY: Boston
;       STATE: MA
;       COUNTRY: US
;       ZIP: 02110-2804
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Diskette
;       COMPUTER: IBM Compatible
;       OPERATING SYSTEM: DOS
;       SOFTWARE: FastSEQ Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/10/096,241
;       FILING DATE: 12-Mar-2002
;       CLASSIFICATION: <Unknown>
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 08/699,591
;       FILING DATE: 19-AUG-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Fasse, J. Peter
;       REGISTRATION NUMBER: 32,983
;       REFERENCE/DOCKET NUMBER: 07334/022001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 617-542-5070
;       TELEFAX: 617-542-8906
;       TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 1476 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     MOLECULE TYPE: cDNA
;     FEATURE:
;       NAME/KEY: Coding Sequence
;       LOCATION: 69...1475
;     OTHER INFORMATION:
;     SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-096-241-7

```

```

Query Match          47.7%;  Score 427.8;  DB 14;  Length 1476;
Best Local Similarity 89.8%;  Pred. No. 3.9e-111;
Matches 459;  Conservative 0;  Mismatches 52;  Indels 0;  Gaps 0;

```

```

Qy      363 CCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGACTG 422
        || || |  ||  |  ||  || ||  || ||  ||  |  ||
Db      98  CCGCGGCAAGAAGCACCCAGAGGGGAGGAAGCGGGAGAGGGAGCCCGATCCCGGGGAGAA 157

```

Qy 423 CGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 482  
 |||  
 Db 158 AGCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCA 217  
 Qy 483 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 542  
 |||  
 Db 218 ATCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGA 277  
 Qy 543 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 602  
 |||  
 Db 278 TGGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAGAAAGAA 337  
 Qy 603 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 662  
 |||  
 Db 338 CTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGTCTGCGAGGC 397  
 Qy 663 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 722  
 |||  
 Db 398 CGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAGCGTGAGCAC 457  
 Qy 723 CACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCTATTG 782  
 |||  
 Db 458 CACCCTGTCATCCTGGTTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCTATTG 517  
 Qy 783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842  
 |||  
 Db 518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAAATGTCC 577  
 Qy 843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873  
 |||  
 Db 578 AAATGGATTCTTCGGACAGAGATGTTTGAG 608

RESULT 5

US-10-096-241-31

; Sequence 31, Application US/10096241

; Publication No. US20020127594A1

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES  
 ; AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/096,241



```

Db      458 CACCCTGTCATCCTGGTCGGGGCACGCCCCGGAAGTGCAACGAGACAGCCAAGTCCTATTG 517
QY      783 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAGTGTCC 842
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      518 CGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGCAAATGTCC 577
QY      843 TGTGGGATACACCGGGGACAGGTGTCAGCAG 873
        |||| | ||| | || ||| | ||
Db      578 AAATGGATTCTTCGGACAGAGATGTTTGGAG 608

```

RESULT 6

```

US-10-271-416-6
; Sequence 6, Application US/10271416
; Publication No. US20040043021A1
; GENERAL INFORMATION:
; APPLICANT: Keith, Tim
; APPLICANT: Little, Randall D.
; APPLICANT: Van Eerdewegh, Paul
; APPLICANT: Dupuis, Josee
; APPLICANT: Del Mastro, Richard G.
; APPLICANT: Allen, Kristina
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO RESPIRATORY DISEASES AND OBESITY
; FILE REFERENCE: 2976-4045
; CURRENT APPLICATION NUMBER: US/10/271,416
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: 60/328,424
; PRIOR FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 22693
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-271-416-6

```

```

Query Match          47.4%; Score 424.8; DB 13; Length 22693;
Best Local Similarity 94.2%; Pred. No. 5.3e-110;
Matches 452; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

```

```

QY      1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      20809 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 20868
QY      61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      20869 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 20928
QY      121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      20929 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 20988
QY      181 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      20989 CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG 21048

```

Qy 241 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 300  
 |||  
 Db 21049 GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG 21108  
 Qy 301 CGCTACATCTTTTCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 360  
 |||  
 Db 21109 CGCTACATCTTTTCTGGAGCCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC 21168  
 Qy 361 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 420  
 |||  
 Db 21169 CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC 21228  
 Qy 421 TGC-GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAA 479  
 ||| | | ||| | | ||| | ||| ||| |||  
 Db 21229 TGCGGTGAGTCGCCCCCTCCCTTTGCTGGAGAAAGGGGGGAGGGGCGAGGTGGTGGAGAA 21288

# RESULT 7

US-10-447-839A-10

; Sequence 10, Application US/10447839A

; Publication No. US20040018181A1

; GENERAL INFORMATION:

; APPLICANT: Kufe, Donald W.

; APPLICANT: Kharbanda, Surender

; APPLICANT: Weitman, Steven D.

; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED THEREFROM

; FILE REFERENCE: 1000.05.009

; CURRENT APPLICATION NUMBER: US/10/447,839A

; CURRENT FILING DATE: 2003-05-29

; PRIOR APPLICATION NUMBER: 10/293,391

; PRIOR FILING DATE: 2002-11-13

; PRIOR APPLICATION NUMBER: 09/951,938

; PRIOR FILING DATE: 2001-09-11

; PRIOR APPLICATION NUMBER: 60/231,841

; PRIOR FILING DATE: 2000-09-11

; NUMBER OF SEQ ID NOS: 109

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 10

; LENGTH: 1054

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-447-839A-10

Query Match 47.3%; Score 424; DB 16; Length 1054;

Best Local Similarity 100.0%; Pred. No. 4.3e-110;

Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 60  
 |||  
 Db 589 ATGAGGCGCGACCCGGCCCCCGGCTTCTCCATGCTGCTCTTCGGTGTGTGCTCGCTCGCCTGC 648  
 Qy 61 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 120  
 |||  
 Db 649 TACTCGCCCAGCCTCAAGTCAGTGCAGGACCAGGCGTACAAGGCACCCGTGGTGGTGGAG 708  
 Qy 121 GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG 180

Db	709		GGCAAGGTACAGGGGCTGGTCCCAGCCGGCGGCTCCAGCTCCAACAGCACCCGAGAGCCG	768
Qy	181		CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	240
Db	769		CCCGCCTCGGGTCGGGTGGCGTTGGTAAAGGTGCTGGACAAGTGGCCGCTCCGGAGCGGG	828
Qy	241		GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	300
Db	829		GGGCTGCAGCGCGAGCAGGTGATCAGCGTGGGCTCCTGTGTGCCGCTCGAAAGGAACCAG	888
Qy	301		CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	360
Db	889		CGCTACATCTTTTTCTGGAGCCACGGAACAGCCCTTAGTCTTTAAGACGGCCTTTGCC	948
Qy	361		CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	420
Db	949		CCCCTCGATACCAACGGCAAAAATCTCAAGAAAGAGGTGGGCAAGATCCTGTGCACTGAC	1008
Qy	421		TGCG	424
Db	1009		TGCG	1012

RESULT 8

US-10-096-241-3

; Sequence 3, Application US/10096241

; Publication No. US20020127594A1

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; Busfield, Samantha J.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; AND USES THEREFOR

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/096,241

; FILING DATE: 12-Mar-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/699,591

; FILING DATE: 19-AUG-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Fasse, J. Peter

; REGISTRATION NUMBER: 32,983

; REFERENCE/DOCKET NUMBER: 07334/022001





RESULT 9  
 US-10-096-241-1  
 ; Sequence 1, Application US/10096241  
 ; Publication No. US20020127594A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gearing, David P.  
 ; Busfield, Samantha J.  
 ; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES  
 ; AND USES THEREFOR  
 ; NUMBER OF SEQUENCES: 33  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson P.C.  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: US  
 ; ZIP: 02110-2804  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSEQ Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/096,241  
 ; FILING DATE: 12-Mar-2002  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/699,591  
 ; FILING DATE: 19-AUG-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Fasse, J. Peter  
 ; REGISTRATION NUMBER: 32,983  
 ; REFERENCE/DOCKET NUMBER: 07334/022001  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617-542-5070  
 ; TELEFAX: 617-542-8906  
 ; TELEX: <Unknown>  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2467 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: circular  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: Coding Sequence  
 ; LOCATION: 79...1893  
 ; OTHER INFORMATION:  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-10-096-241-1

Query Match 45.2%; Score 405.4; DB 14; Length 2467;  
 Best Local Similarity 87.9%; Pred. No. 1e-104;  
 Matches 442; Conservative 0; Mismatches 61; Indels 0; Gaps 0;



```
; FEATURE:
; OTHER INFORMATION: MAP TO CHR5.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.55
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.72
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.66
; OTHER INFORMATION: SWISSPROT HIT: O14511, EVALUE 3.00e-29
; OTHER INFORMATION: NT HIT: AF119152.1, EVALUE 1.00e-109
; OTHER INFORMATION: EST_HUMAN HIT: BF108794.1, EVALUE 3.00e-93
US-10-029-386-26613
```

```
Query Match          19.3%; Score 173; DB 15; Length 201;
Best Local Similarity 100.0%; Pred. No. 5.1e-39;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      424 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 483
          |||
Db      27  GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 86

Qy      484 TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 543
          |||
Db      87  TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 146

Qy      544 GGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAG 596
          |||
Db      147 GGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAG 199
```

# RESULT 11

```
US-10-447-839A-11
; Sequence 11, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Surender
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED
THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 419
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-447-839A-11
```

```
Query Match          19.3%; Score 173; DB 16; Length 419;
Best Local Similarity 100.0%; Pred. No. 6.1e-39;
```

```

Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      424 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 483
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       50  GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 109

Qy      484 TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 543
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      110 TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 169

Qy      544 GGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAG 596
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      170 GGCAAGGAGCTCAACCGCAGCCGAGACATTCGCATCAAATATGGCAACGGCAG 222

```

RESULT 12

US-10-029-386-12913

; Sequence 12913, Application US/10029386

; Publication No. US20030194704A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES  
USEFUL FOR GENE

; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

; FILE REFERENCE: AEOMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029,386

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 34288

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 12913

; LENGTH: 573

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO CHR5.3

; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.55

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.72

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.66

; OTHER INFORMATION: SWISSPROT HIT: O14511, EVALUE 2.00e-28

; OTHER INFORMATION: NT HIT: AF119152.1, EVALUE 0.00e+00

; OTHER INFORMATION: EST\_HUMAN HIT: BG996653.1, EVALUE 1.00e-108

US-10-029-386-12913

```

Query Match          19.3%; Score 173; DB 15; Length 573;
Best Local Similarity 100.0%; Pred. No. 6.6e-39;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      424 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 483
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      377 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 436

Qy      484 TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 543
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      437 TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 496

```

Qy 544 GGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAG 596  
 |||  
 Db 497 GGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAG 549

RESULT 13

US-10-271-416-7

; Sequence 7, Application US/10271416  
 ; Publication No. US20040043021A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Keith, Tim  
 ; APPLICANT: Little, Randall D.  
 ; APPLICANT: Van Eerdewegh, Paul  
 ; APPLICANT: Dupuis, Josee  
 ; APPLICANT: Del Mastro, Richard G.  
 ; APPLICANT: Allen, Kristina  
 ; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES  
 ; TITLE OF INVENTION: RELATING TO RESPIRATORY DISEASES AND OBESITY  
 ; FILE REFERENCE: 2976-4045  
 ; CURRENT APPLICATION NUMBER: US/10/271,416  
 ; CURRENT FILING DATE: 2002-10-11  
 ; PRIOR APPLICATION NUMBER: 60/328,424  
 ; PRIOR FILING DATE: 2001-10-11  
 ; NUMBER OF SEQ ID NOS: 9  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 7  
 ; LENGTH: 45450  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(45450)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-10-271-416-7

Query Match 19.3%; Score 173; DB 13; Length 45450;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-38;  
 Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 424 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 483  
 |||  
 Db 7026 GCCACCCGGCCCAAGTTGAAGAAGATGAAGAGCCAGACGGGACAGGTGGGTGAGAAGCAA 7085

Qy 484 TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 543  
 |||  
 Db 7086 TCGCTGAAGTGTGAGGCAGCAGCCGGTAATCCCCAGCCTTCCTACCGTTGGTTCAAGGAT 7145

Qy 544 GGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAG 596  
 |||  
 Db 7146 GGCAAGGAGCTCAACCGCAGCCGAGACATTTCGCATCAAATATGGCAACGGCAG 7198

RESULT 14

US-10-447-839A-12

; Sequence 12, Application US/10447839A  
 ; Publication No. US20040018181A1

```

; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Surender
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED
THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/231,841
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-447-839A-12

```

```

Query Match          13.9%; Score 124.6; DB 16; Length 493;
Best Local Similarity 93.5%; Pred. No. 3.5e-25;
Matches 130; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

Qy      594 CAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGT 653
          |||
Db      227 CAGAAAGAACTCACGACTACAGTTCAACAAGGTGAAGGTGGAGGACGCTGGGGAGTATGT 286
          |||

Qy      654 CTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAG 713
          |||
Db      287 CTGCGAGGCCGAGAACATCCTGGGGAAGGACACCGTCCGGGGCCGGCTTTACGTCAACAG 346
          |||

Qy      714 CGTGAGCACCACCCTGTCA 732
          || || || | ||
Db      347 CGGTAGGTGGGCCAGACA 365

```

# RESULT 15

```

US-10-447-839A-13
; Sequence 13, Application US/10447839A
; Publication No. US20040018181A1
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald W.
; APPLICANT: Kharbanda, Surender
; APPLICANT: Weitman, Steven D.
; TITLE OF INVENTION: MUC1 INTERFERENCE RNA COMPOSITIONS AND METHODS DERIVED
THEREFROM
; FILE REFERENCE: 1000.05.009
; CURRENT APPLICATION NUMBER: US/10/447,839A
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 10/293,391
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 09/951,938
; PRIOR FILING DATE: 2001-09-11

```

; PRIOR APPLICATION NUMBER: 60/231,841  
; PRIOR FILING DATE: 2000-09-11  
; NUMBER OF SEQ ID NOS: 109  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 13  
; LENGTH: 350  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-447-839A-13

Query Match 13.8%; Score 124; DB 16; Length 350;  
Best Local Similarity 100.0%; Pred. No. 4.7e-25;  
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      715 GTGAGCACCACCCTGTCATCCTGGTCGGGGGCACGCCC GGAAGTGCAACGAGACAGCCAAG 774
          |||
Db      99 GTGAGCACCACCCTGTCATCCTGGTCGGGGGCACGCCC GGAAGTGCAACGAGACAGCCAAG 158

Qy      775 TCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGC 834
          |||
Db     159 TCCTATTGCGTCAATGGAGGCGTCTGCTACTACATCGAGGGCATCAACCAGCTCTCCTGC 218

Qy      835 AAGT 838
          |||
Db     219 AAGT 222
```

Search completed: August 15, 2004, 12:23:39  
Job time : 491.839 secs